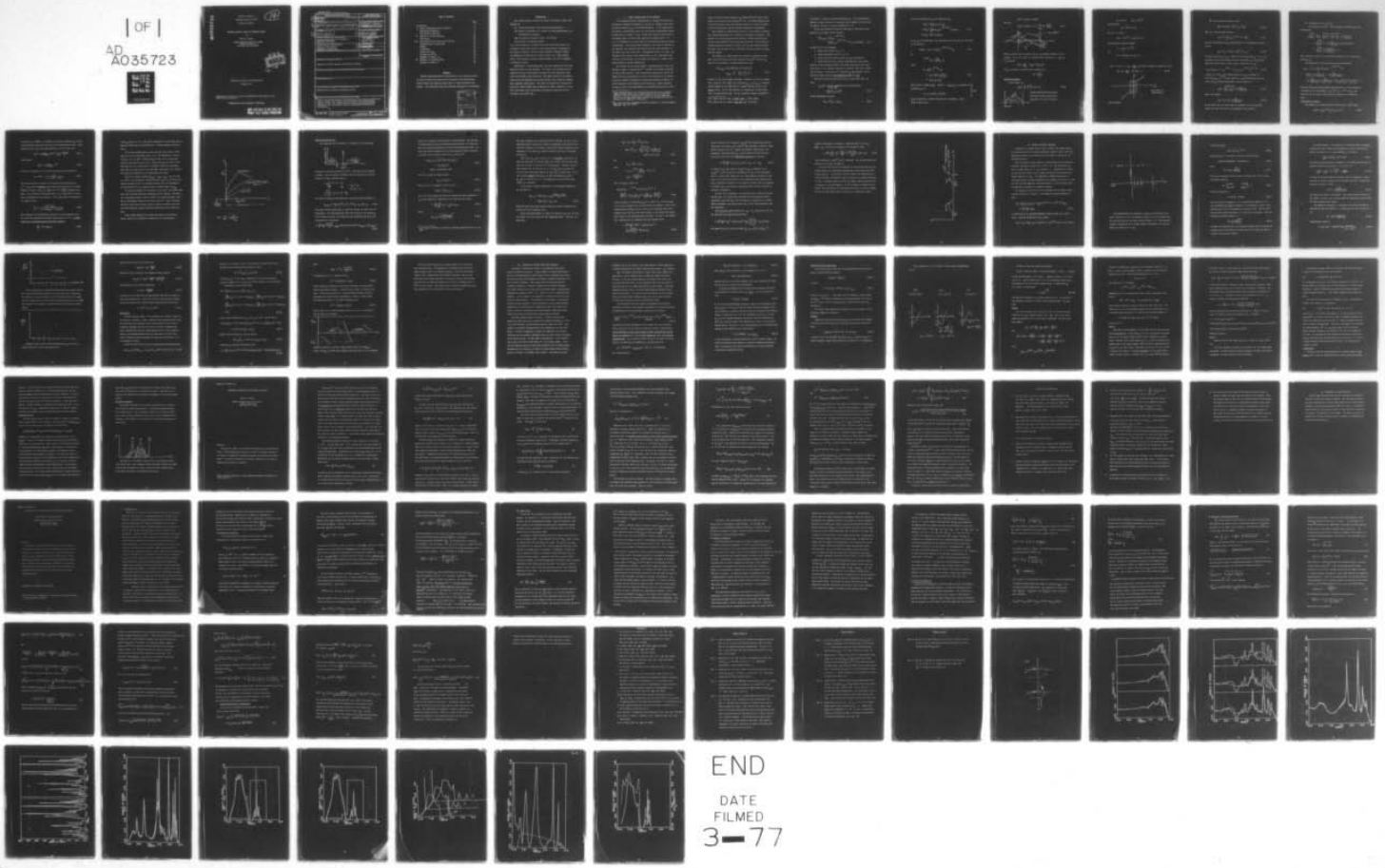
AD-A035 723 BELFER GRADUATE SCHOOL OF SCIENCE NEW YORK F/G 20/12
ELECTRONIC ENERGY LEVELS IN IMPERFECT SOLIDS.(U)
JAN 77 D C MATTIS N00014-76-C-0690
UNCLASSIFIED TR-1 NL

| OF |
AD A035723



END

DATE FILMED
3-77

TECHNICAL REPORT #1

CONTRACT N00014-76-C-0690

Project NR 392-015

ELECTRONIC ENERGY LEVELS IN IMPERFECT SOLIDS

by

Daniel C. Mattis

BELFER GRADUATE SCHOOL OF SCIENCE
Yeshiva University
New York, New York 10033

Prepared for Office of Naval Research

January 1977

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: TR-1	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): ELECTRONIC ENERGY LEVELS IN IMPERFECT SOLIDS	5. TYPE OF REPORT & PERIOD COVERED: Interim technical report
6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s): Daniel C. Mattis	8. CONTRACT OR GRANT NUMBER(s): N00014-76-C-0690
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Belfer Graduate School of Science, Yeshiva University, New York, N.Y. 10033	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 89 p.
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Department of the Navy, Arlington, Virginia 22217	12. REPORT DATE: January 1977
	13. NUMBER OF PAGES: 86
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report)
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Resease; Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Imperfections, Surfaces, Electrons, Solids

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Several scattering-theoretic formulations of the theory of solids are used to discuss the energy levels associated with imperfections, surface states, and electron energy levels in highly disordered substances. The theoretical bases are thoroughly developed and illustrated.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

UNCLASSIFIED 053 820

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

TABLE OF CONTENTS

Abstract

Several scattering-theoretic formulations of the theory of solids are used to discuss the energy levels associated with imperfections, surface states, and electron energy levels in highly disordered substances. The theoretical bases are thoroughly developed and illustrated.

INTRODUCTION

This Progress Report concerns the theory of imperfect solids, with emphasis on:

(I) atomic imperfections (and their optical properties),

(II) effects of proximity to a surface on these imperfections (e.g. "recombination centers"),

(III) the nature of surfaces proper, and finally,

(IV) transport in imperfect solids.

It is in the nature of a technical foray into solid state theory, to illustrate current applications of scattering-theoretic techniques to the study of vacancies, electron states in highly disordered solids, the mathematical formulation of surface states, etc. Other applications: small polaron, excitons, impurity bands, etc. will be examined in subsequent reports.

Additionally, a forthcoming report will deal specifically with more practical aspects of this research: the study of the photoelectric decomposition of H_2O ("photolysis") by means of n-type TiO_2-based solar cells or by similar other materials. The topics treated in the present report have been inspired by the study of these "more practical" aspects. Other immediate applications: the study of catalysis and corrosion begins with a model foreign atom or molecule or defect proximate to a surface; such models can be analyzed by the methods illustrated in Part III below, as we shall show.

I. ATOMIC IMPERFECTIONS IN AN INSULATOR

In the Wannier-function representation, a foreign atom at site R_o introduces a diagonal perturbation V_o as well as a change in the overlap matrix elements K_{oj} to the nearest-neighbors. If the magnitude of the former is sufficiently great, or of the latter sufficiently reduced, a bound-state is formed.* If $V_o > 0$ the state comes out of the top of the valence band, the bound state known as an acceptor-level. If $V_o < 0$ it condenses out of the bottom of the conduction band and creates a donor-level. If the point-group symmetry of the solid is lowered by the impurity, the conduction and valence bands may both contribute to the bound state level and create a "recombination-center" or trap. The symmetry may be lowered either spontaneously, or by a Jahn-Teller distortion, or by proximity to the surface (cf. Chapter II, below), each having different optical consequences.

The vacancy may be treated (albeit, unconventionally) by the same methods as apply to impurity atom.** Consider the F-center (a missing Cl atom in NaCl crystal): while conventional studies have created the bound state out of Na orbitals on atoms in the vicinity of the vacancy, we find it simpler to postulate the continued existence of Wannier orbitals centered at the position of the vacancy R_o, shifted upward by an amount V_o due to the modification of the Madelung potential at that

---

*A long-range potential such as the Coulomb potential has an infinite number of bound states; the occupancy being by one or at most two electrons, the extra states are not usually of crucial importance. Thus we use the zero-range model, in which the perturbation V_i is localized at the impurity, in this work.

**This point has been independently noted by others: cf. Jaros and Brand, Phys. Rev. B14, 4494 (1976).

site, and having overlap integrals K_{oj} modified from the values appropriate to the levels of an ordinary Cl atom. The vacancy Wannier-orbital will be 1s-like rather than 3p-like (there is no core of occupied states at the vacancy, hence no exclusion-principle requirements).

The O vacancy in reduced TiO_2 is equally a color center, although more complicated because it is required to accomodate 2 electrons. The problem of 2 interacting electrons orbiting a common attractive well has been solved exactly,* and it has been found that if the 2-body repulsive potential U exceeds a critical value U_c both electrons cannot be bound. But even when $U < U_c$, the first electron is easier to ionize than the second.

Thus taking as a simple model of an impurity or vacancy the one first studied by Slater and Koster some 20 years ago** ($V_o \neq 0$, K_{oj} same as in the absence of the imperfection) we have:

$$(H')_{ni,mj} \equiv \int d_3 r \, \phi_n^*(r-R_i) \, H'(r) \, \phi_m(r-R_j)$$

$$= g_{n,m} \cdot \begin{cases} 0 & \text{for } i \text{ or } j \neq o \\ 1 & \text{if } i = j = o \end{cases} , \qquad \text{(I.1)}$$

because H', ϕ_i, ϕ_j are highly localized. Moreover, if $H'(r)$ is reasonably constant in the atomic cell at $\vec{R}_o$ then $g_{n,m} \sim g_n \, \delta_{n,m}$, is approximately diagonal in the band index (a constant potential doesn't mix atomic levels). (We'll also consider the consequences of band mixing which are very interesting). Let $g_n \equiv$ coupling constant (strength of

*D.C. Mattis and E.H. Lieb, J. Math. Phys. 7, 2045 (1966).

**G.F. Koster and J.C. Slater, Phys. Rev. 95, 1167 (1954).

perturbing — potential, previously denoted V_o). If a substituional impurity is more attractive to electrons than a regular host atom that it replaces, then $g < 0$; if less attractive, $g > 0$.

We can use definition of Wannier functions to find matrix representation in the Bloch states (nk,mk'):

$$(H')_{nk,mk'} = \frac{1}{N} g_{nm} e^{i(k'-k)\cdot R_o}$$

$$= \frac{1}{N} g_n \delta_{n,m} e^{i(k'-k)\cdot R_o} \quad \text{for } g_{n,m} \text{ diagonal.} \qquad (I.2)$$

We shall now do the following:

1. Obtain bound states of $H_o + H'$
2. Obtain scattering states of $H_o + H'$
3. Obtain Scattering Cross Section of these ("mobility")
4. Obtain Optical matrix elements connecting the bound states with any of the conduction states, and obtain resulting optical spectrum, that is, do a complete analysis of "acceptor" ($g > 0$) or "donor" ($g < 0$) levels. At the end, we'll consider g_{nm} <u>not</u> diagonal, for an exactly solvable theory of <u>recombination centers</u> or <u>traps</u>.

For now, let's start with bound state (if any) which takes the form:

$$\Phi_{n,o}(\vec{r}) = \text{linear combination of all Bloch states in nth band}$$

$$= \frac{1}{\sqrt{N}} \sum_k F_{n,k} \psi_{nk}(\vec{r}) . \qquad (I.3)$$

We use Schrodinger equation:

$$(H_o + H')\Phi_{no} = E_{no}\Phi_{no} \qquad (I.4)$$

to get the coefficients F_{nk} and eigenvalue E_{no}

$$H_o\Phi_{no} = \frac{1}{\sqrt{N}} \sum_k \epsilon_{nk} F_{nk} \psi_{nk}$$

$$+ H'\Phi_{no} = \frac{1}{N^{3/2}} g_n \sum_{kk'} e^{i(k'-k)\cdot R_o} F_{nk'}\psi_{nk'} \qquad \text{(I.5)}$$

$$= E_{no}\Phi_{no} = \frac{1}{\sqrt{N}} \Sigma F_{nk} E_{no} \psi_{nk}$$

As the ψ_{nk} are orthonormal, the coefficients must be equal on both sides of the equation.

$$F_{nk}(E_{no} - \epsilon_{nk}) = g_n \frac{1}{N} \sum_{k'} e^{i(k'-k)\cdot R_o} F_{nk'} \qquad \text{(I.6)}$$

So:

$$e^{ik\cdot R_o} F_{nk} = \frac{g_n}{E_{no} - \epsilon_{nk}} \Delta_n$$

where

$$\Delta_n \equiv \frac{1}{N} \sum_{k'} e^{ik'\cdot R_o} F_{nk'}$$

$$= - g_n \Delta_n \frac{1}{N} \sum_{k'} \frac{1}{\epsilon_{nk'} - E_{no}} \qquad \text{(I.7)}$$

$$\equiv - g_n \Delta_n S_n (E_{no})$$

which serves to define $S_n(E)$. Either $1 = -g_n S_n(E_{no})$ is satisfied (determining E_{no}) or

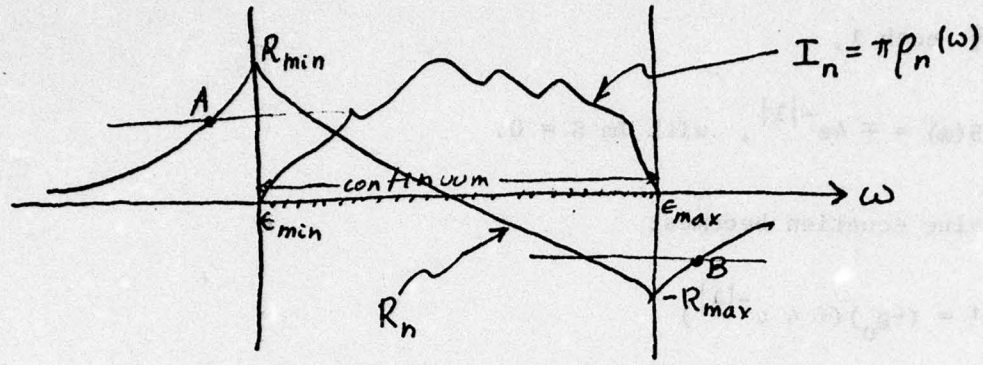
$$\Delta_n = 0 \text{ (trivial solution)} \qquad \text{(I.8)}$$

$S_n(\omega)$ is generally a complex function of its argument. Using $\text{Im}\,\frac{1}{X} = \pi\delta(X)$ we have:

$$I_n(\omega) \equiv \mathrm{Im}S_n(\omega) = \pi\rho_n(\omega)$$

and also

$$R_n(\omega) \equiv \mathrm{Re}S_n(\omega) = \text{p.p.} \int_{\epsilon_{min}}^{\epsilon_{max}} d\omega' \, \frac{\rho_n(\omega')}{\omega' - \omega} \qquad \text{(I.9)}$$



Thus Δ_n equation has no solution for E_n in continuum, because r-h-s is complex. If $g_n < 0$, there is a solution below continuum, $1 = -g_n R_n$ at "A" provided

$$|g_n| > \frac{1}{R_{min}} \quad (R_{min} \equiv R_n(\epsilon_{min}))$$

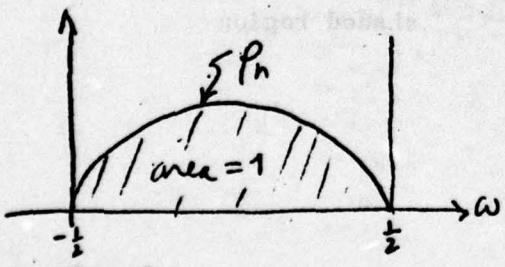
If $g_n > 0$ there's a solution above continuum if

$$g_n > \left|\frac{1}{R_{max}}\right|$$

<u>Analytical Example</u>:

To fix ideas, let

$$\rho_n(\omega) = \frac{8}{\pi}\sqrt{\tfrac{1}{4} - \omega^2}, \quad |\omega| < \tfrac{1}{2} \qquad \text{(I.10)}$$



proper behavior near band edges

but omits van Hove singularities

Bandwidth = 1 = unit of energy.

Then for

$$\omega \equiv \tfrac{1}{2} \cos\theta, \qquad S(\omega) = -4e^{-i\theta},$$

and we verify

$$I_n = 4 \sin\theta = \pi\rho_n$$

For $\omega = \pm \tfrac{1}{2} \cosh\lambda$,

$$S(\omega) = \mp 4e^{-|\lambda|}, \quad \text{with Im } S = 0.$$

The eigenvalue equation becomes:

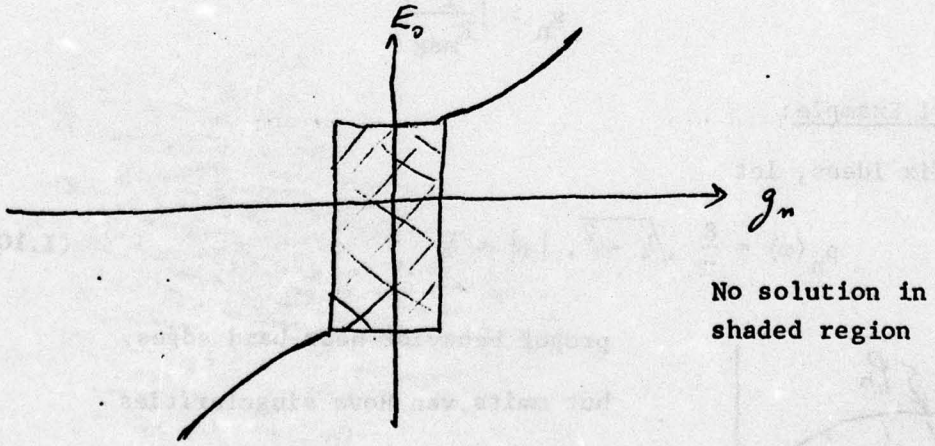
$$1 = (-g_n)(\mp 4\, e^{-|\lambda|})$$

or

$$|\lambda| = \ln(\pm 4g_n) \quad \text{for } |g_n| > \tfrac{1}{4}.$$

$E_0 = \pm \tfrac{1}{2} \cosh\lambda = \pm \left(g_n = \frac{1}{16g_n}\right)$, *and after throwing out unphysical root,*

$$E_0 = g_n + \frac{1}{16g_n} \quad \text{for } |g_n| > \tfrac{1}{4}. \qquad \text{(I.11)}$$



No solution in shaded region

End of Example.

For the continuum solutions we take

$$\Phi_{nk} = A_{n,k}\{\psi_{nk} + \frac{1}{N}\sum_{k'} L_{k,k'}\,\psi_{nk'}\} \tag{I.12}$$

where A_k = normalization constant,

$$A_k = \{1 + \frac{1}{N^2}\,|L_{k,k'}|^2\}^{-\frac{1}{2}} = 1 + O(\tfrac{1}{N}) \tag{I.13}$$

Equating coefficients of ψ_{nk} on both sides of the Schrodinger equation yields:

$$E_{nk} = \epsilon_{nk} + N^{-1}(g_n + g_n N^{-1} \sum_{k'} L_{k,k'}) \tag{I.14}$$

For many purposes $E_{nk} = \epsilon_{nk} + O(1/N)$ can be replaced by ϵ_{nk} — i.e. the continuum of *exact scattering solutions* "interlaces" the continuum of unperturbed Bloch states. Next, equate coefficients of $\psi_{nk'}$:

$$L_{k,k'}(E_{nk} - \epsilon_{nk'}) = g_n(e^{ik\cdot R_o} + N^{-1}\sum_{k''} L_{k,k''}\, e^{ik''\cdot R_o})e^{-ik'\cdot R_o} \tag{I.15}$$

A series of obvious steps follows (solve for $L_{k,k'}$ in terms of $D \equiv N^{-1}\sum_{k''} L_{k,k''}\, e^{ik''\cdot R_o}$ then solve for D self-consistently using the value of $L_{k,k''}$ obtained previously) easily leading to:

$$L_{k,k'} = \frac{e^{i(k-k')\cdot R_o}}{\epsilon_{nk} - \epsilon_{nk'}}\;\frac{g_n}{1 + g_n S_n(\epsilon_{nk})} \tag{I.16}$$

Then, (14) becomes:

$$E_{nk} = \epsilon_{nk} + \frac{1}{N}\;\frac{g_n}{1 + g_n S_n(\epsilon_{nk})} \tag{I.17}$$

We now verify that the bound state is orthogonal to the scattering states, and that the latter are orthogonal to one another.

(a) Orthogonality of Φ_{no} to Φ_{nk}:

The integral of (I.3)* X (I.13) yields: (recalling ψ_{nk}'s are orthonormal):

$$\frac{e^{ik\cdot R_o}}{\epsilon_k - E_o} - \frac{g_n}{N}\sum_{k'} \frac{e^{ik'\cdot R_o}}{\epsilon_{k'} - E_o} \frac{e^{i(k-k')\cdot R_o}}{\epsilon_{k'} - \epsilon_k} \frac{1}{1 + g_n S_n(\epsilon_{nk})}$$

$$= \frac{e^{ik\cdot R_o}}{\epsilon_k - E_o}\left\{1 - \frac{g_n}{1 + g_n S_n(\epsilon_{nk})}\left[S_n(\epsilon_{nk}) - S_n(E_o)\right]\right\}$$

$$= \frac{e^{ik\cdot R_o}}{\epsilon_k - E_o}\left\{\frac{1 + g_n S_n(E_o)}{1 + g_n S_n(\epsilon_{nk})}\right\} = 0$$

This vanishes by virtue of the eigenvalue equations (I.7), (I.8) which determine E_{no}.

(b) Orthogonality of one scattering state with another:

$$L_{k,k'} + L^*_{k'k} + \frac{1}{N}\sum_{k''} L^*_{k'k''} L_{kk''} = e^{i(k-k')\cdot R_o} \times$$

$$\left\{\frac{1}{\epsilon_k - \epsilon_{k'}}\left[\frac{g_n}{1 + g_n S_n(\epsilon_{nk})} - \frac{g_n}{1 + g_n S_n^*(\epsilon_{nk'})}\right]\right.$$

$$\left. + \frac{g_n}{1 + g_n S_n(\epsilon_{nk})} \frac{g_n}{1 + g_n S_n^*(\epsilon_{nk'})} \cdot \frac{1}{N}\sum_{k''} \left(\frac{1}{\epsilon_{nk''} - \epsilon_{nk'}}\right)^* \left(\frac{1}{\epsilon_{nk''} - \epsilon_{nk}}\right)\right\} = 0$$

assuming appropriate infinitesimal inaginary parts in each denominator (according to the prescription: $\epsilon_k + i0^+$, $\epsilon_{k'} - i0^+$ for $L_{k,k'}$ and similarly for the others, satisfying the requirements of causality thus).

Scattering by Impurity

The differential scattering cross-section $\sigma_k(\theta) = |f_k|^2$, where

$$f_k(\theta) = \lim_{r \to \infty} rN^{-1} \sum_{k'} L_{k,k'} \psi_{k'}(r) \qquad \text{(I.18)}$$

in which $\cos\theta = \vec{k}\cdot\vec{k}'/kk'$. We neglect the periodic component u_{nk} of the Bloch functions, and perform the sum over the phase factors only. After integrating over angles and approximating ϵ_k by k^2 we have left:

$$f_k(\theta) \sim \frac{g}{1 + gS(\epsilon_k)} \int dk' k'^2 \frac{1}{k'^2 - ?} \frac{\sin k'r}{k'} \qquad \text{(I.19)}$$

which yields:

$$f_k(\theta) \sim \frac{g}{1 + gS(\epsilon_k)} e^{ikr} \qquad \text{(I.20)}$$

As this is independent of θ the total scattering cross-section is:

$$\sigma_k = \text{const. } x \frac{g^2}{(1 + g R)^2 + (g\pi\rho)^2} \qquad \text{(I.21)}$$

The scattering time is related to the scattering cross-section by $\tau_k^{-1} = v_k \sigma_k$, *and the* <u>*mobility*</u> *(the conductivity per carrier at a given* energy) is $e^2\tau_k/m^*$, where $v_k = \overline{\nabla_k \epsilon_k}$ and $1/m^* = \overline{\nabla_k \nabla_k \epsilon_k}$ (suitably averaged). For a parabolic band approximation, $v_k \propto \rho(\epsilon_k)$ and $m^* = \text{const.}$ This is sufficient for our purposes. For N_I impurities (concentration $c_I = N_I/N$) we have:

$$\frac{1}{\tau_k} = c_I \frac{2\pi g^2 \rho(\epsilon_k)}{(1 + gR(\epsilon_n))^2 + (g\pi\rho(\epsilon_k))^2} \qquad \text{(I.22)}$$

after juggling the multiplicative constants to obtain agreement with the first Born approximation (Fermi's Golden Rule) at $g = 0$. <u>An identical result is obtained directly from the relation</u>:

$$\frac{1}{2\tau_k} = N_I \, \text{Im}\{E_{n,k}\} \qquad \text{(I.23)}$$

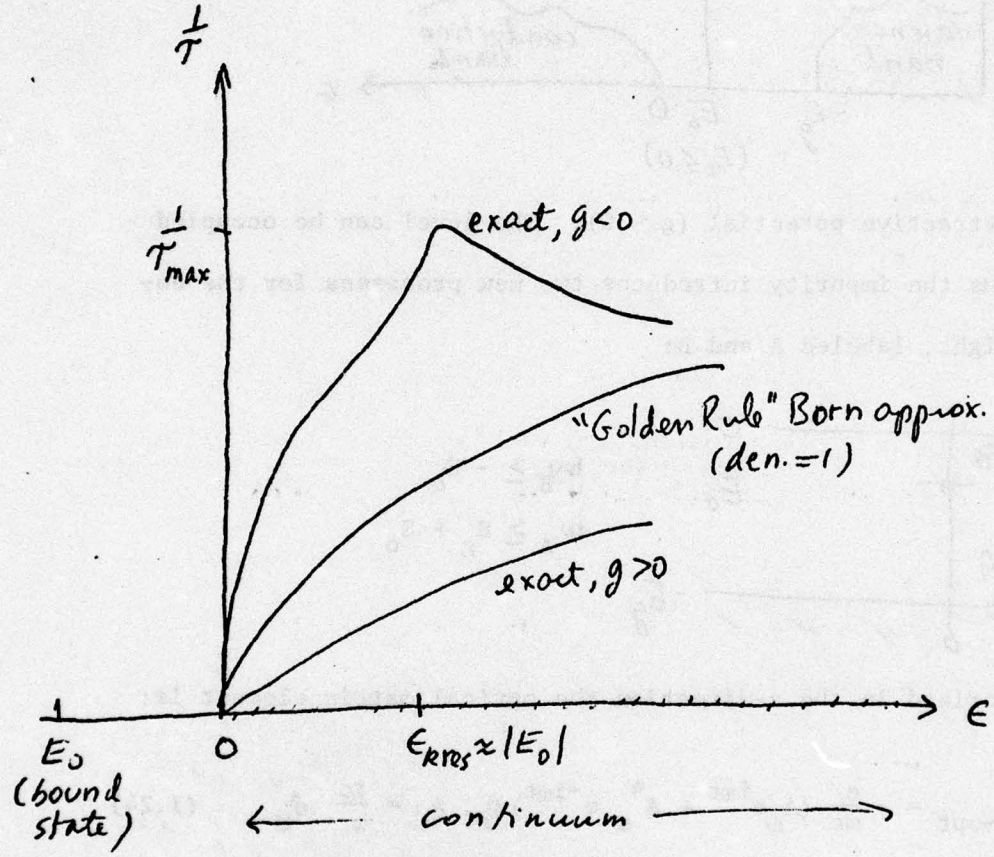
with E_{nk} given in (I.17). The general dependence of scattering time on energy is indicated in the figure below. Several remarks are also in order:

(1) the Born approximation is very bad when there exists a bound state, even at the band edge, where $\rho \to 0$. The denominator, instead of being 1 is $(1 + gR)^2 = g^2(R(\epsilon_k) - R(E_o))^2$ which can be either much larger or much smaller than 1; for a bound state, E_o, very close to the band edge (as for a typical donor or acceptor level), this expression will tend to be very small at the band edge, resulting in true scattering being much enhanced over the approximate Born value.

(2) At the resonance energy ($\epsilon_{k_{res.}} - \epsilon_{min} \approx \epsilon_{min} - E_o$) the term $(1 + gR)^2$ in the denominator vanishes entirely, and the scattering time becomes extremely short; the coupling constant g^2 in the denominator and the factor g^2 in the numerator cancel. We have there $(1/\tau)_{max}$.

(3) When the potential is repulsive so that there is no bound state below the conduction band and no resonance at low energies, the effective scattering strength is $g^2/(1 + gR)^2$ at the lower band edge, which may be considerably reduced from the Born value g^2. Thus a repulsive potential which may be strong enough to create an acceptor level near the valence band may induce very little scattering in the conduction band.

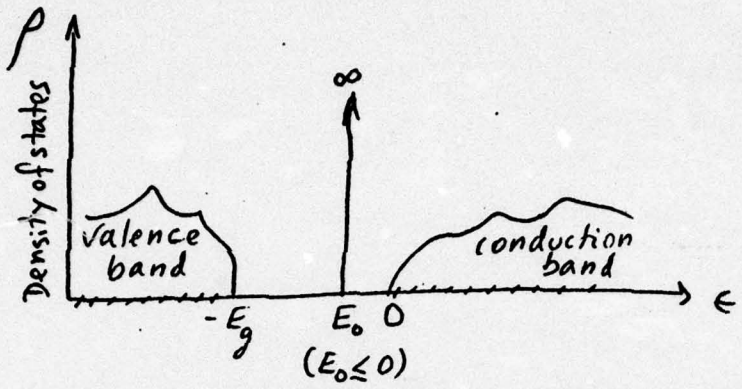
These remarks summarize some important aspects of scattering theory, which can be explicitly verified for our soluble model.



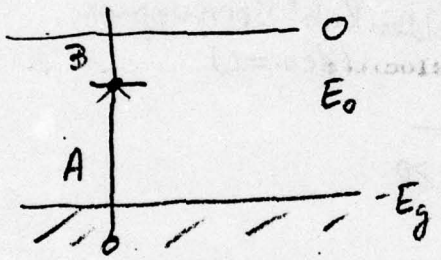
Note: $\frac{1}{\tau_{max}} = \frac{2c_I}{\pi\rho(\epsilon_{k_{res}})}$

Some Optical Properties:

The energetics for an insulator + 1 impurity is the following:



assuming an attractive potential ($g < 0$). The level can be occupied or empty: thus the impurity introduces two new processes for the absorption of light, labeled A and B:



For light polarized in the z-direction the optical matrix element is:

$$H_{el\text{-}opt} = -\frac{e}{mc}\left(A_\omega e^{i\omega t} + A^*_\omega e^{-i\omega t}\right) p_z, \quad A_\omega = \frac{ic}{\omega}\,\delta_\omega \qquad (I.24)$$

This weak coupling Hamiltonian is currectly treated in lowest Born approximation. For long wavelength light (not X-rays) we may assume A_ω to be constant in space, and obtain for the probability per unit time of process A or B occuring:

$$w = \frac{2\pi}{\hbar}\left(\frac{e}{mc}\right)^2\left(\frac{|\delta_\omega|^2 c^2}{\omega^2}\right) \sum_{final} |(final|p_z|initial)|^2\, \delta(\epsilon_{final} \pm \hbar\omega - \epsilon_{final}) \qquad (I.25)$$

with (+) for emission of photon and (-) for absorption. For the sake of definiteness we treat the absorption spectrum only, the reader can make the necessary modifications to obtain the emission spectrum. Because the spatial dependence of the electromagnetic field is ignored, we require only the matrix elements of p_z, which are:*

$$(p_z)_{n'k',nk} \equiv \int d_3 r \; \psi^+_{n'k'}(r) p_z \; \psi_{nk}(r)$$
$$= \delta_{k,k'} M_{n'n}(k) \tag{I.26}$$

(Note: conservation of $\vec{k}$)

For $n'=n$, we have the simple result:

$$M_{n,n} = m v_z(n,k) \tag{I.27}$$

where v_z is the z-component of the velocity:

$$\vec{v}(n,k) = \frac{1}{\hbar} \vec{\nabla}_k \epsilon_{n,k} \tag{I.28}$$

For $n' \neq n$ there exists an f-sum rule relating the matrix elements of p_z to one compoent (zz-component) of the effective mass tensor:

$$1 - \frac{m}{\hbar^2} \frac{\partial^2}{\partial k_z^2} \epsilon_{n,k} = \sum_{n' \neq n} f_{n',n;z} \tag{I.29}$$

where

$$f_{n',n;z} \equiv \frac{2}{m} \frac{|(p_z)_{n'k,nk}|^2}{\epsilon_{n'k} - \epsilon_{nk}} \tag{I.30}$$

---

*A.H. Wilson, "The Theory of Metals", Cambridge University Press, p. 46, 1954.

The other components and a proof are given in Wilson, op. cit. In an approximate model where only 2 bands are considered, this sum rule relates the curvature of the bands to the optical matrix element; so the parameters in the model are closely related and are not to be chosen arbitrarily.

Now, either $M_{n',n}(k)$ is zero at $k = 0$ (forbidden transition, e.g. from an s-like band to a d-like band, both of which have the same parity), but it won't be zero at $k \neq 0$: thus $M_{n'n}(k) = \text{constant} \times k_z$ is a good estimate for its value at $k = 0$. Or, $M_{n',n}(0)$ is non-zero at $k = 0$, as for transitions between an s-like and a p-like band. In the case of such allowed transitions, we shall approximate $M_{n'n}(k) \cong M_{n',n}(0)$ by its vale at $k = 0$, for simplicity, and treat the matrix element as a constant.

In the study of optical properties, it is frequently helpful to use the identity

$$|M_{n',n}| = \left|\int d_3r\ \psi^+_{n'k}\ z\ \psi_{nk}\right| \left|\frac{m(\epsilon_{n'k} - \epsilon_{nk})}{\hbar}\right|$$

$$= \left|\int d_3r\ \psi^+_{n'k}\ z\ \psi_{nk}\right|\ m|\omega| \qquad \text{(I.31)}$$

where the last result uses energy conservation, and is therefore only correct for use in equation (I.25).

We can now study process A. Label the initial states (v) for "valence band", the final states (c) for "conduction band". They are, respectively:

$$\phi_{vk} = \psi_{vk} + \frac{1}{N} \sum_{k'} L^{(v)}_{k,k'} \psi_{vk'}$$

$$\text{where } L^{(v)}_{k,k'} = \frac{e^{i(k-k')\cdot R_o}}{\epsilon_{vk} - \epsilon_{vk'}} \frac{g_v}{1-g_v S_v(\epsilon_{vk})}$$

(Note: see (I.16)) (I.32)

and

$$\phi_{c,o} = \frac{1}{\sqrt{N}} \sum F_{ck} \psi_{ck} \tag{I.33}$$

where

$$F_{ck} = e^{-ik\cdot R_o} \frac{g_c \Delta_c}{E_o - \epsilon_{ck}}, \text{ (by (I.7)}$$

$$g_c \Delta_c = \left(\frac{\partial S_c(E_o)}{\partial E_o}\right)^{-\frac{1}{2}}$$

Thus, the matrix element is:

$$(\phi_{vk}|p_z|\phi_{co}) = e^{-ik\cdot R_o} (\partial S_c(E_o)/\partial E_o)^{-\frac{1}{2}} N^{-\frac{1}{2}} X$$

$$\left\{\frac{M_{vc}(k)}{E_o - \epsilon_{ck}} + \frac{1}{N} \sum_{k'} \left(\frac{1}{\epsilon_{vk} - \epsilon_{vk'}}\right)^* \frac{g_v}{1-g_v S_v^*(\epsilon_{vk})} \frac{M_{cv}(k')}{E_o - \epsilon_{vk'}}\right\} \tag{I.34}$$

If the transition is forbidden, $M_{cv}(k')$ is odd in k'_z and the second term in curly brackets (the sum) will vanish by symmetry. Even if the transitions are allowed, for $g < 0$ (as we are assuming) there is no resonance near the top of the valence band, so the second term contributes little to the absorption near threshold. We drop it for simplicity, obtaining for the transition rate per impurity atom:

$$w = \frac{2\pi}{\hbar} \left(\frac{e}{m\omega}\right)^2 |\mathcal{M}|^2 \times (\partial S_c(\epsilon_o)/\partial E_o)^{-1} \times$$

$$\frac{1}{N} \sum_k \left(\frac{M_{cv}(k)}{E_o - \epsilon_{ck}}\right)^2 \delta(E_o - \hbar\omega - \epsilon_{vk}) \tag{I.35}$$

which is readily seen to behave as $(\omega-\omega_T)^{\frac{1}{2}}$ near threshold ω_T when the transitions are allowed, and $(\omega-\omega_T)^{3/2}$ when forbidden, rising to a peak before falling off as ω^{-4}. Before this happens, absorption at the fundamental absorption edge $\hbar\omega_g - E_g$ has already obscured the line. From (I.25) and (I.26) the <u>fundamental absorption</u> is given by

$$w = \frac{2\pi}{\hbar}\left(\frac{e}{m\omega}\right)^2 |\mathcal{E}_\omega|^2 \sum_k |M_{cv}(k)|^2 \, \delta(\epsilon_{ck} - \hbar\omega - \epsilon_{vk}) \qquad \text{(I.36)}$$

It starts at E_g, also rises as $(\omega-\omega_g)^{\frac{1}{2}}$ if M_{cv} is constant, or as $(\omega-\omega_g)^{3/2}$ if the transition is forbidden, but the overall absorption rate is a factor N/N_I larger than the total impurity absorption (every site in the crystal contributes to the fundamental edge, whereas only the N_I impurities contribute to the impurity process; thus the impurity is visible only in the range of frequencies where the pure crystal is otherwise transparent).

We next turn to process B; if the bound level is close to the conduction band, the optical line associated with this process will be reasonably narrow and quite low in frequency as compared to the fundamental absorption, thus this process will be more distinguishable from the background.

The absorption matrix element is: $M_{nn} = mv_z$ (equation (I.4)), so the absorption rate <u>per</u> impurity becomes:

$$w = \frac{2\pi}{\hbar}\left(\frac{e}{m\omega}\right)^2 |\mathcal{E}_\omega|^2 \left(\partial S_c(E_o)/\partial E_o\right)^{-1} \frac{1}{N} \sum_k \frac{m^2 v_z^{\,2}(k)}{(E_o-\epsilon_{ck})^2} \, \delta(\epsilon_{ck}-\hbar\omega-E_o)$$

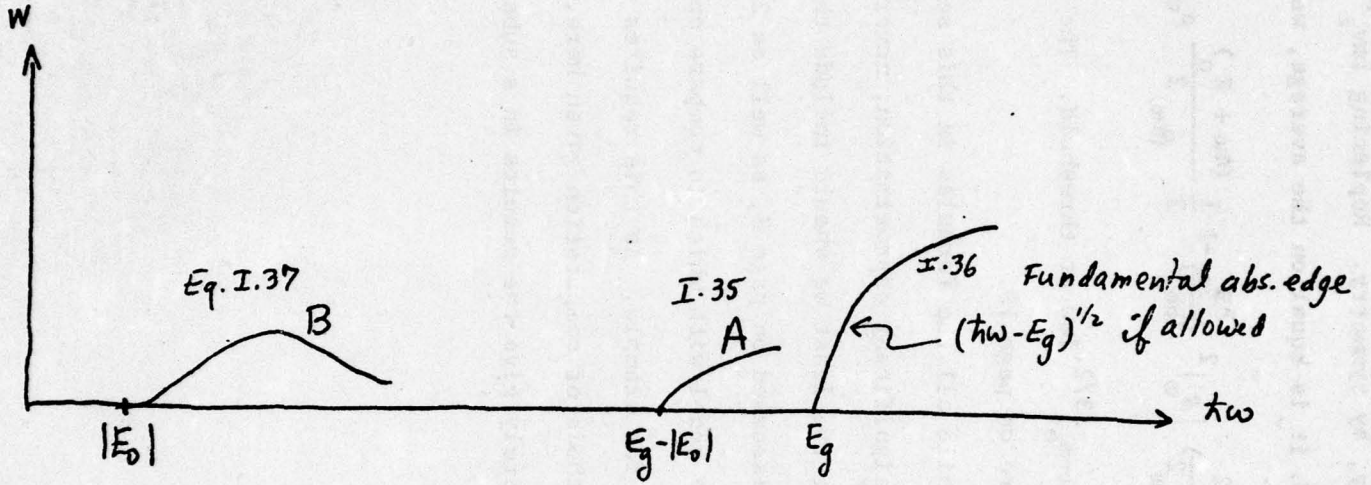
The summed term in the matrix element, $\sum_{k'} L_{k,k'} v_z(k') \, (E_o-\epsilon_{ck'})^{-1}$

vanishes identically, by symmetry. Replacing Imv_z^2 by $1/3\ \epsilon_{ck} = \frac{1}{3}(\hbar\omega + E_o)$, to which it is equal on the average, we find:

$$w = \frac{2\pi}{\hbar}\left(\frac{e^2}{m\omega^2}\right)\left|\mathcal{R}_\omega\right|^2\left(\frac{\partial S_c}{\partial E_o}\right)^{-1}\frac{1}{3}\frac{(\hbar\omega + E_o)}{(\hbar\omega)^2}\rho_c(\hbar\omega + E_o) \qquad (I.37)$$

also satisfying a $(\omega-\omega_o)^{3/2}$ law at threshold. The various results are plotted in the figure on page 19.

We note that while all the formulas in this section have been obtained without any significant approximation, nevertheless the model is not accurate. At the least we should include the modification in overlap integrals discussed on page 3, as well as 2-electron repulsion to have a reasonable model with which to compare optieal properties of 0 vacancies in TiO_2 for example. As this requires no fundamental departure from the methods of computation given here, we shall no expand on this but shall merely give the results in a subsequent report.


W
Eq. I.37
B
I.35
A
I.36
Fundamental abs. edge
$(\hbar\omega - E_g)^{1/2}$ if allowed
$\hbar\omega$
$|E_0|$
$E_g - |E_0|$
E_g

II. EFFECTS OF SURFACE PROXIMITY

Proximity to a surface will be seen to affect the binding energy, to cause structure in the optical absorption spectrum, and to lower the point-group symmetry of the impurity site to where it can act as a recombination center.

For the purpose of this chapter, we shall assume that the boundary condition typifying a surface is $\psi = 0$ at $z = 0$. The solid will consist of atomic planes at $z = na_o$, $n=1,2,3,\ldots$. An impurity at a depth $d = pa_o$ from the surface can be rendered soluble by the artifice of an "image" impurity at $-pa_o$. The eigenstates, whether bound-state or scattering states, are then computed for the infinite solid ($-\infty < n < +\infty$), but only functions <u>odd</u> under reflection about $z = 0$ are retained, thus automatically satisfying the boundary condition.

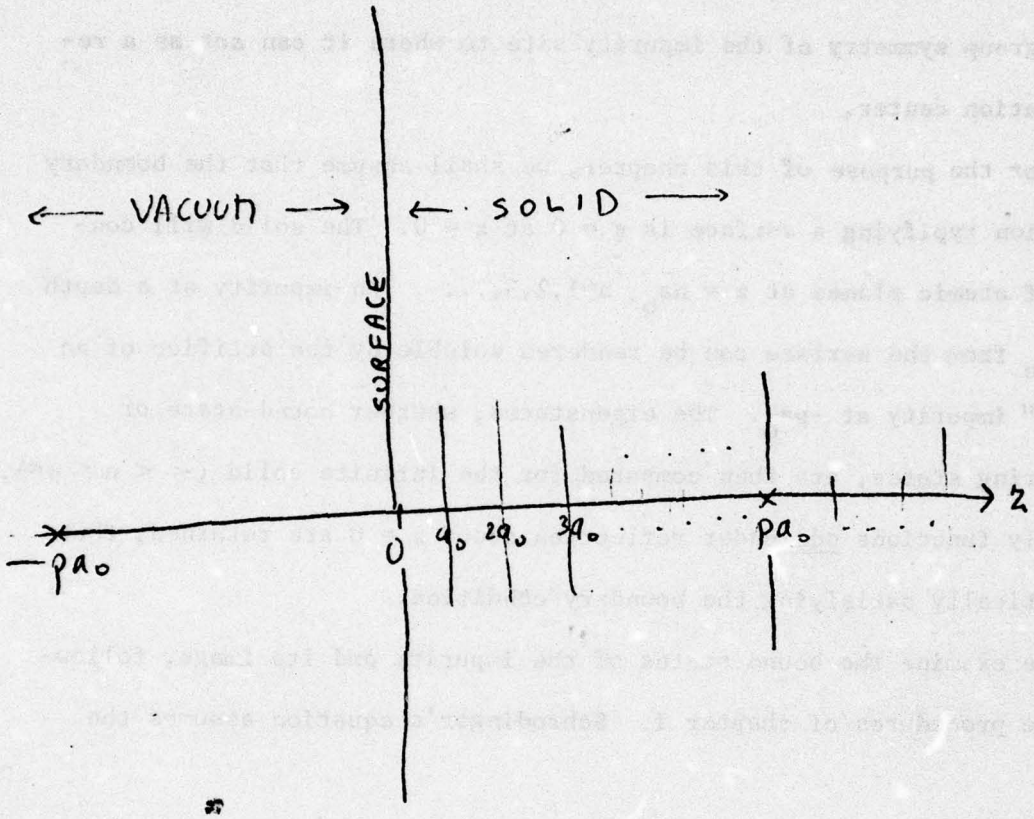
We examine the bound states of the impurity and its image, following the procedures of chapter I. Schrodinger's equation assumes the form:

$$H\Phi_o = \frac{1}{\sqrt{N}} \sum_k F_k \left\{\epsilon_k \psi_k(r) - \frac{g}{N} \sum_{k'} (e^{i(k'-k)\cdot d} + e^{-i)k'-k)\cdot d}) \psi_{k'}\right\}$$

$$= -\Delta(d) \frac{1}{\sqrt{N}} \Sigma F_k \psi_k \qquad \text{(II.1)}$$

in which $\Delta(d)$ is the distance-dependent binding energy (i.e. $\Delta(\infty) = |E_o|$). Equating coefficients of ψ_k yields:

$$F_k(\Delta + \epsilon_k) = \frac{g}{N} \sum_{k'} F_{k'} \, 2 \cos(k'-k)\cdot d = g(e^{-ik\cdot d} \, \xi(d) + c.c.) \qquad \text{(II.2)}$$



All eigenfunctions are required to vanish on the surface of the solid, defined by $z = 0$. We extend the solid to $z < 0$ but retain only the solutions that are odd under inversion about $z = 0$. This procedure requires introduction of an "image" impurity potential at $z = -pa_o$ to balance an impurity at $z = pa_o$.

in which we define

$$\xi(d) = \frac{1}{N} \Sigma F_{k'} e^{ik' \cdot d} \tag{II.3}$$

The self-consistent solution of (II.2) and (II.3) yields:

$$\xi(\pm d) = g[(\xi(\pm d)T(0) + \xi(\mp d)T(\pm 2d))] \tag{II.4}$$

where

$$T(x) \equiv \frac{1}{N} \Sigma_{k'} \frac{e^{ik' \cdot x}}{\Delta(d)+\epsilon_{k'}} \tag{II.5}$$

The pair of equations (II.4) can have a solution only if the secular determinant vanishes:

$$\text{Det} \begin{Vmatrix} 1-gT(0) & -gT(2d) \\ -gT(-2d) & 1-gT(0) \end{Vmatrix} = 0 \tag{II.6}$$

which is equivalent to

$$1 = g(T(0) \pm T(2d)) \tag{II.7}$$

Only the solution belonging to (-) is physically acceptable, as it alone satisfies the boundary condition; it corresponds to $\xi(d) = -\xi(-d)$ and thus $D_k = 0$ at $d = 0$ according to equation (II.2). The (+) solution is even at about the plane $z = 0$, and must be discarded. Thus the eigenvalue equation reduces to:

$$\frac{1}{g} = \frac{1}{N} \Sigma_k \frac{1-e^{2ik \cdot d}}{\Delta(d)+\epsilon_k} \tag{II.8}$$

We shall now establish that it is entirely possible for g to satisfy the condition for the existence of a bound state in the bulk, yet fail to satisfy (II.8) near the surface.

For this purpose, it is sufficient to calculate $g_c(d)$, the minimum g which will yield a bound state according to equation (II.8) above. Deep in the bulk, at $d = \infty$, upon setting $\Delta_o = 0$ we have:

$$\frac{1}{g_c(\infty)} = \frac{1}{N} \Sigma \frac{1}{\epsilon_k} = S(0) = R(0) \tag{II.9}$$

with $S(\omega)$ previously defined in chapter I. The change upon nearing the surface, formed by setting $\Delta(d) = 0$ in (II.8), is:

$$\frac{1}{g_c(\infty)} - \frac{1}{g_c(d)} = \frac{1}{N} \Sigma \frac{e^{2ik\cdot d}}{\epsilon_k} = \frac{2m}{4\pi|2d|} \tag{II.10}$$

in which m = effective mass of electrons, and the integrand is recognized as the Fourier transform of the Coulomb potential. Solving this equation for $g_c(d)$, we have:

$$g_c(d) = \frac{g_c(\infty)}{1-(mg_c(\infty)/4\pi|d|} \tag{II.11}$$

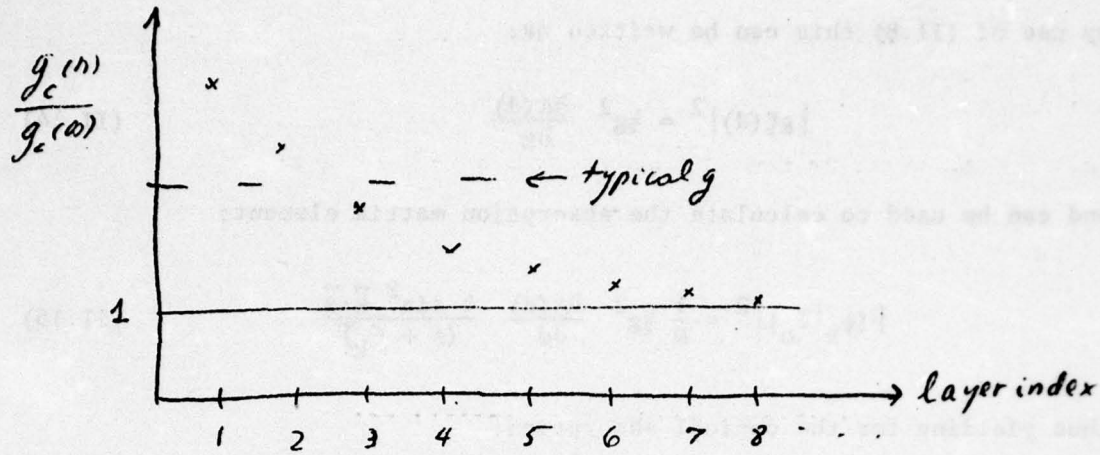
The results are shown in the figures. The binding energy decreases as the surface is neared, and may even vanish for impurities located within the first several atomic layers.

The effects of proximity to a surface on the optical absorption spectrum are even more dramatic. We obtain them by first solving for D_k.

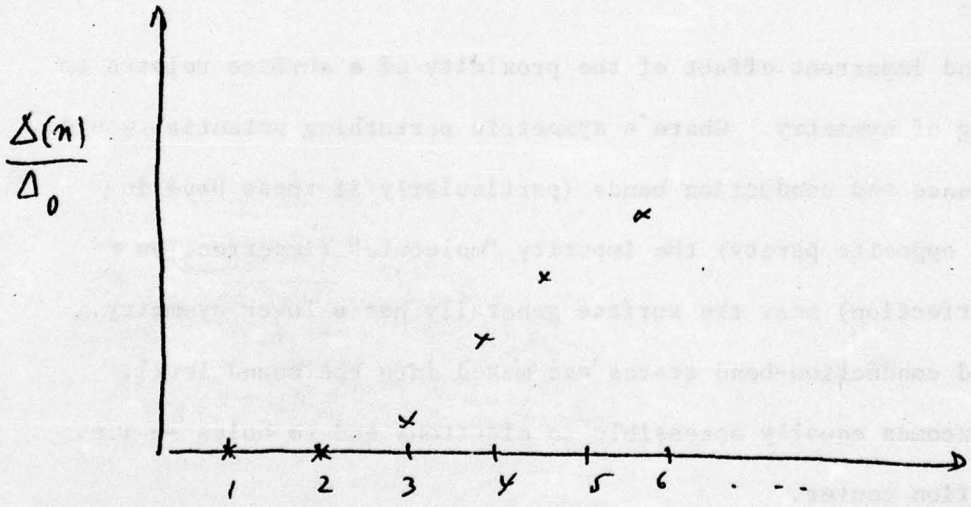
$$F_k = g\xi(d) \frac{2i \sin k\cdot d}{\Delta(d) + \epsilon_k} \tag{II.12}$$

Normalization requires:

$$1 = \frac{1}{N} \Sigma |F_k|^2 = 2g^2|\xi(d)|^2 \frac{1}{N} \Sigma \frac{1-\cos 2k\cdot d}{(\epsilon_k + \Delta(d))^2} \tag{II.13}$$



Plot of $g_c(n)/g_c(\infty)$ as schematically obtained from equation (II.11). For a typical attractive potential g a bound state may be produced in the bulk but not in the surface; for the illustrated value, the criterion for a bound state is met only for the surface layers $n \geq 3$, no bound states exist if the impurity is within either of the first two atomic planes.



Binding energy in units of bulk binding-energy $\Delta_o \equiv \Delta(\infty)$ for a typical potential, such as illustrated above.

By use of (II.8) this can be written as:

$$|g\xi(d)|^2 = \tfrac{1}{2}g^2 \frac{\partial\Delta(d)}{\partial g} \tag{II.14}$$

and can be used to calculate the absorption matrix element:

$$|(\psi_k|\Phi_o)|^2 = \frac{1}{N}\tfrac{1}{2}g^2 \frac{\partial\Delta(d)}{\partial d} \frac{4\sin^2 \vec{k}\cdot\vec{d}}{(\Delta + \epsilon_k)} \tag{II.15}$$

thus yielding for the optical absorption:

$$w = w_\infty\left(1 - \frac{\sin 2kd}{2kd}\right) \tag{II.16}$$

in which w_∞ is the transition probability/unit time given in chapter I, with $\Delta(d)$ replacing $\Delta(\infty)$. The extra structure comes from the trigonometric term, which in the effective-mass approximation has the value

$$k = (2m)^{\frac{1}{2}}(\omega - E_g + \Delta(d))^{\frac{1}{2}}. \tag{II.17}$$

Band Mixing:

A second important effect of the proximity of a surface relates to the lowering of symmetry. Where a symmetric perturbing potential would not mix valence and conduction bands (particularly if these have intrinsically opposite parity) the impurity "molecule" (imperfection + mirror imperfection) near the surface generally has a lower symmetry. Valence- and conduction-band states are mixed into the bound level, which now becomes equally accessible to electrons and to holes — i.e. a recombination center.

A model for this starts with the following perturbing Hamiltonian:

$$H'_{n'k',nk} = \frac{2}{N}\{g_n\delta_{n,n'} + g_{n',n}(1-\delta_{n,n'})\}\sin k\cdot R_o \sin k'\cdot R_o \tag{II.18}$$

Restrict n to 2 bands "c" and "v" and obtain the bound state if any. The bound state wavefunction must take the form:

$$\Phi_o = N^{-\frac{1}{2}} \sum_{nk} F_{n,k}\, \psi_{n,k}(r-R_o) \qquad \text{(II.19)}$$

in which $\psi_{n,k}(r)$ is the linear combination of two Bloch functions, (k_x,k_y,k_z) <u>minus</u> that at $(k_x,k_y,-k_z)$, which vanishes on the z =0 plane.

Schrodinger's equation now reads:

$$H\Phi_o = \frac{1}{\sqrt{N}} \sum_k F_{vk}\, \epsilon_{vk}\, \psi_{vk} + \frac{1}{\sqrt{N}} \Sigma\, F_{ck}\, \epsilon_{ck}\, \psi_{ck}$$

$$+ \frac{2g_v}{N^{3/2}} \sum_{kk'} \sin k\cdot R_o \sin k'\cdot R_o F_{vk} \psi_{vk'} + \frac{2g_{cv}}{N^{3/2}} \sum_{kk'} \sin k\cdot R_o \sin k'\cdot R_o F_{vk} \psi_{ck'}$$

$$+ \frac{2g_c}{N^{3/2}} \sum_{kk'} \sin k\cdot R_o \sin k'\cdot R_o F_{ck} \psi_{ck'} + \frac{2g_{vc}}{N^{3/2}} \sum_{kk'} \sin k\cdot R_o \sin k'\cdot R_o F_{ck} \psi_{vk'}$$

$$= E_o \Phi_o \qquad \text{(II.20)}$$

and we equate coefficients of ψ_{vk} and ψ_{ck} in turn. For the first:

$$(E_o - \epsilon_{ck}) F_{vk} = 2 \sin k\cdot R_o \left[g_v \frac{1}{N} \sum_{k'} F_{vk'} \sin k'\cdot R_o + g_{vc} \frac{1}{N} \sum_{k'} F_{ck'} \sin k'\cdot R_o \right]$$

$$\equiv 2 \sin k\cdot R_o \left[g_v D_v + g_{vc} D_c \right] \qquad \text{(II.21)}$$

which serves to define D_n. For the second:

$$(E_o - \epsilon_{ck}) F_{ck} = 2 \sin k\cdot R_o \left[g_{cv} D_v + g_c D_c \right] \qquad \text{(II.22)}$$

Evaluating D_n using the definitions above:

$$D_v = \frac{1}{N} \sum_{k'} (E_o - \epsilon_{ck'})^{-1} \{ (2 \sin^2 k'\cdot R_o)(g_v D_v + g_{vc} D_c) \} = -S_v(E_o)\left[g_v D_v + g_{vc} D_c \right] \qquad \text{(II.23)}$$

where

$$S_n(E) = N^{-1} \sum_k \frac{2 \sin^2 k \cdot R_o}{\epsilon_n(k) - E} \qquad \text{(II.24)}$$

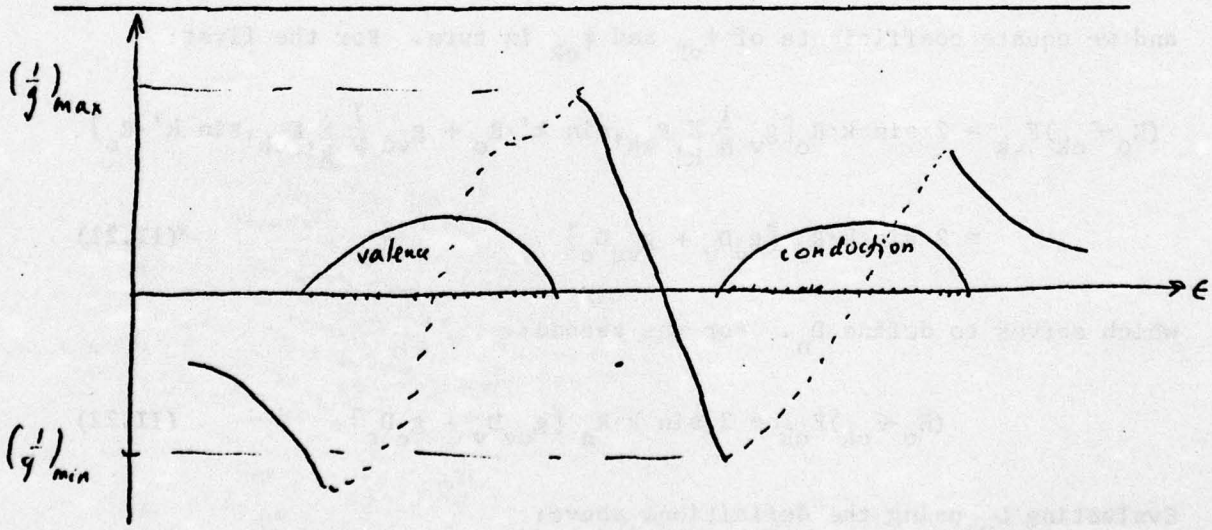
by analogy with (I.7). Similarly for D_c:

$$D_c = -S_c(E_o)[g_c D_c + g_{cv} D_v] \qquad \text{(II.25)}$$

These homogeneous equations (II.23) and (II.25) have a solution iff a secular determinant vanishes. A special case of this is trivially soluble: let all g's be equal (i.e. strong interband mixing), then the condition for the bound state (recombination center) to exist within the energy gap reduces to the simple algebraic form:

$$g^{-1} = -[S_v(E_o) + S_c(E_o)] \qquad \text{(II.26)}$$

We plot the r.h.s. of this equation as well as the d.o.s. functions of the two bands in the Figure:



Graphical analysis of (II.26): shows solution exists if $(1/g)_{max} > (1/g) > (1/g)_{min}$ (hatched line represent real part when r.h.s. is complex).

The bound state evidently has a dipole moment if the two bands have opposite parity. The magnitude of the dipole moment can be obtained through (I.31) as a function of $M_{vc}(k)$. We omit this calculation, but note: a first-order Stark effect is implied. That is, the bound state will respond linearly to an applied electric field $\vec{\mathcal{E}}$ while the bulk state, in which the bands are not mixed, is affected only to $0(\vec{\mathcal{E}}^2)$ in such a field. This should help to identify recombination centers optically, by the techniques of electro-reflectance.

III. CONCERNING ELECTRON STATES NEAR SURFACES

The study of surfaces has become a very important and active branch of solid state physics. Surface physics is particularly important in the understanding of solid state devices, including solar cells, also for the study of heterogeneous catalysis, including photolysis, all for a variety of reasons. The methods by which surfaces have been studied are typically: numerical computation of a desired quantity (e.g. band structure) for a thin slab of N atomic planes, varying N and expanding the property as $XN^0 + YN^1$. Clearly, X is the surface-related property. Just as clearly, this procedure is tedious and does not lend itself to rapid insights. We illustrate in this chapter a scattering-theoretic method designed to focus specifically on the surface-related properties in a solid where N is immediately taken to be large or infinite. By way of illustration we shall determine, under what conditions will electronic surface state energy bands be created.

Our model problem starts as follows: we find the linear combination of Bloch functions to satisfy the boundary conditions $\psi = 0$ at $z \leq 0$, and evidently, the Schrodinger equation within the solid. This analysis is extremely illuminating and has, apparently, not been done before. We find it important to analyze the amplitude of the eigenfunctions near the surface, especially near the energy minima and maxima of the band structure. We then apply a perturbation: in the simplest case, only on the first atomic plane (i.e. the surface plane). If we ask the following question: under what conditions will an infinitesimal perturbation ($g \ll$ either E_g or width of energy bands) produce surface states?, we obtain a remarkably simple answer: only when, in the 2D

Brillouin zone of the surface, there exist points or lines along which a certain component of the inverse-effective-mass tensor, α_{zz}, vanishes.

When the surface perturbation is large, there always exists surface states, and our method permits a relatively simple evaluation thereof. In any case, having obtained the surface bound states and the scattering bulk states of the terminated solid, we can proceed to study the effects of an added perturbation at the surface. (This perturbation could be done to a molecule undergoing heterogeneous catalysis or photolysis.) We shall see that the effects of such a perturbation are qualitatively different than if, as in the usual analysis, the bulk wave functions and band structures were used.

In the periodic solid the complete orthonormal set of Wannier functions is related to the complete orthonormal set of Bloch functions (eigenfunctions of the unperturbed, periodic H_o) as follows:

$$\underset{\text{(Bloch)}}{\psi_{nk}(r)} = N^{-\frac{1}{2}} \sum_j e^{ik\cdot R_j} \phi_n(r-R_j) \text{ and } \underset{\text{(Wannier}}{\phi_n(r-R_j)} = N^{-\frac{1}{2}} \sum_k e^{-ik\cdot R_j} \psi_{nk}(r) \qquad \text{(III.1)}$$

Thus they are Fourier transforms of one another over the N discrete values of R_i or k (in the space or reciprocal lattice, respectively). The introduction of one or more surfaces and/or perturbations has the effect of replacing $\exp \pm ik\cdot R_i$ by the components $U_k(R_i)$ of a unitary transformation. As we intend to retain the ϕ_n as the basis, it is important to reformulate the operator H_o, initially given as

$$H_o = -(\hbar^2/2m)\nabla^2 + V_{\text{periodic}}(r) \quad \text{(with } \hbar = 1 \text{ henceforth)}.$$

One readily obtains:

$$\int d_3 r\ \phi_n^*(r-R_j) H_o f(r) = \epsilon_n(-i\partial/\partial\vec{R}_j) F(R_j) \qquad \text{(III.2)}$$

where $F(R_j)$ is the coefficient in the expansion of f, i.e.:

$$f(r) \equiv \sum_j F(R_j)\phi_n(r-R_j) \qquad \text{(III.3)}$$

Equation (III.2) is valid for arbitrary f(r) (not necessarily an eigenfunction of H_o or of any other Hamiltonian).

To introduce diagonal (in the Wannier-representation) potentials as perturbations, we merely add: $\Sigma_m V_m \delta_{\vec{R}_m,\vec{R}_j}$ to H_o and finally obtain:

$$H\ U_k(R_j) = E_k U_k(R_j) \qquad \text{(III.4)}$$

as our new, pseudo-Schrodinger equation, subject to the boundary conditions $U_k(R_j) = 0$ for $Z_j \leq 0$ (for the study of surface phenomena; note that periodic b.c. could be used if we were not concerned with surfaces; indeed, we shall use p.b.c. for the coordinates X_j and Y_j), in which E_k is the new energy eigenvalue (which may or may not interlace the old) k is a quantum number which adiabatically reduces to the crystal omentum when $H \to H_o$ and p.b.c. are restored, $U_k(R_j)$ is the "wavefunction" defined on the N points R_j only, and finally:

$$H = \epsilon_n(-i\partial/\partial\vec{R}_j) + \sum_m V_m \delta_{\vec{R}_m,\vec{R}_j} \qquad \text{(III.5)}$$

is the Hamiltonian, a matrix defined only on the N lattice points. It should be emphasized that despite the explicit differential operator, this H incorporates only discrete translations by integer multiples of primitive translation vectors.

Surface-Theoretic Hamiltonian:

For surfaces which retain the bulk translation vectors in the X-Y plane the above simplify somewhat:

$$H\, U_k(Z_j) = E_k U_k(Z_j) \qquad \text{(III.6)}$$

in which

$$H = \epsilon_n(k_x, k_y, -i\partial/\partial Z_j) + \sum_m V_m\, \delta_{Z_m, Z_j} \qquad \text{(III.7)}$$

and $Z_m = ma$, $m=1,2,3,\ldots$ Note that we have a different "linear chain" having $N_z = N^{1/3}$ distinct eigenvectors and eigenfunctions, for each of the $N^{2/3}$ values of the 2D vector (k_x, k_y). This vector we denote $k_{\|}$ henceforth.

The linear chain problem thus defined is far from trivial, as the following example will indicate.

Example:

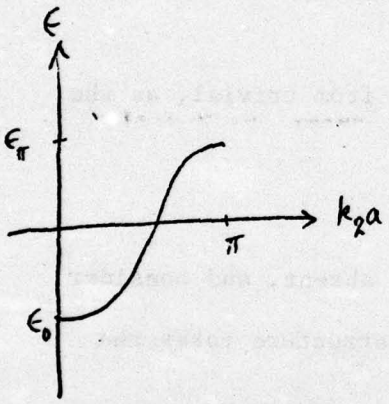
Suppose the potential perturbations V_m to be absent, and consider the solutions of (III.6 and III.7) when the band structure takes the form:

$$\epsilon_n(k_{\|}, k_z) = -2K \cos k_z a - 2L \cos 2k_z a \qquad \text{(III.8)}$$

in which K and L are presumably functions of $k_{\|}$. A plot of $\epsilon_n(k_z)$ is given below ($k_{\|}$ is fixed; only positive k_z are shown as ϵ is symmetric).

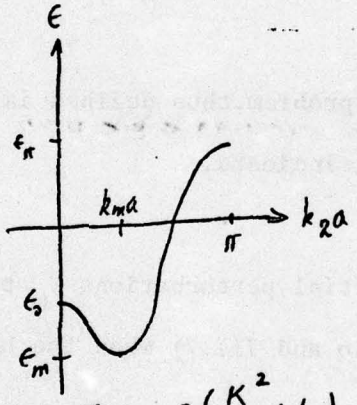
We may always pick $K > 0$; consider various signs and magnitudes for L:

(a) $0 < |L| < \frac{1}{4}K$



$\epsilon_0 = -2(K+L)$, $\epsilon_\pi = 2(K-L)$

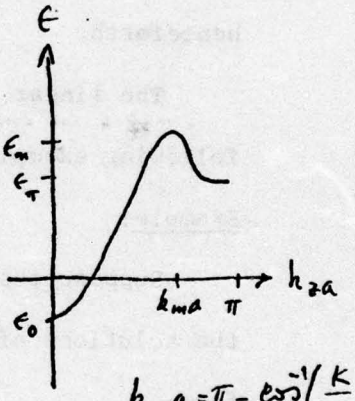
(b) $L < -\frac{1}{4}K < 0$



$\epsilon_m = 2\left(\frac{K^2}{8L} + L\right)$

$k_m a = \cos^{-1}\left(\frac{K}{-4L}\right)$

(c) $L > \frac{1}{4}K > 0$



$k_m a = \pi - \cos^{-1}\left(\frac{K}{4L}\right)$

We wish to solve the eigenvalue equation:

$$H\,U(z) = \{-2K \cosh a\partial/\partial z - 2L \cosh 2a\partial/\partial z\}U(z) = E\,U(z) \qquad \text{(III.9)}$$

on the lattice points $z = ma$, $m=1,2,\ldots$ subject to the b.c. $U = 0$ for $m = 0,-1,\ldots$ and $U =$ finite as $m \to +\infty$. We can construct such a solution using the set of k's which satisfy $\epsilon(k_z) = E$; label them k_q, $q = 1,2,3,4$. We then have:

$$U = \sum_q A_q e^{ik_q z} \qquad \text{(III.10)}$$

and need only determine the A_q which satisfy the b.c. The solutions have different character in cases (a)-(c) pictured above. We start with:

Case a:

For E in the range $-2(K + L) \leq E \leq +2(K - L)$ there are 2 real solutions $\pm k_o$ and 2 complex roots $\pm (k_1 + ik_2)$. One of the latter grows exponentially as $m \to +\infty$ and must be discarded. We find, after some algebra:

$$k_o a = \cos^{-1}\left\{\frac{-K}{4L} + \frac{L}{|L|}\left[\left(\frac{K}{4L}\right)^2 + \frac{2L-E}{4L}\right]^{\frac{1}{2}}\right\}$$

and

$$a(k_1 + ik_2) = \pi\left(\frac{L+|L|}{2|L|}\right) + i \ln\left\{\left|\frac{K}{4L}\right| + \left[\left(\frac{K}{4L}\right)^2 + \frac{2L-E}{4L}\right]^{\frac{1}{2}} + \left[\left(\left|\frac{K}{4L}\right| + \left(\left(\frac{K}{4L}\right)^2 + \frac{2L-E}{4L}\right)\right)^2 - 1\right]^{\frac{1}{2}}\right\}$$

Thus:

$$U(m) = Ae^{ik_o ma} + Be^{-ik_o ma} + Ce^{ik_1 ma - k_2 ma}$$

satisfies the difference equation (III.9) everywhere except at m = 2 and 1. However, imposing U(0) = U(-1) = 0 enables us to satisfy the equation there also. The vanishing condition at m = 0 is:

$$A + B + C = 0$$

while that at m = -1 requires:

$$Ae^{-ik_o a} + Be^{+ik_o a} + Ce^{-i(k_1 + ik_2)a} = 0$$

Normalization of the "wavefunction" U for a chain terminating at N_z yields:

$$|A|^2 + |B|^2 = 1/N_z \quad \text{(in the limit } N_z \gg 1/k_2 a)$$

These uniquely determine a single solution for each value of E. The reader may work out the elementary algebra for the coefficients A, B, C and also verify that when L = 0 this solution properly reduces to

$$U = (2/N_z)^{\frac{1}{2}} \sin k_o am, \text{ where } k_o a = \cos^{-1}(-E/2K).$$

We next turn to:

Case b:

This case is "non-standard", in the sense that the energy minimum occurs not at k = 0. In the range $\epsilon_o \leq E \leq \epsilon_\pi$ the analysis is exactly that of case (a), and a unique solution with 3 components is again found. However, for E in the range $\epsilon_m \leq E \leq \epsilon_o$, which encompasses the neighborhood of the energy minimum, there are found 4 real wavevectors $\pm k_o$ and $\pm k_1$ at each energy. Imposition of the two b.c. at m = 0 and -1 reduces this number to 2 (except precisely at $E = \epsilon_m$ where only a single solution survives — howevew, this is a minor technical point).

Let $\epsilon(k_o) = \epsilon(k_1) = E$ (cf. Figure b). As the k's are all real, we use sin and cos functions instead of exponentials, and guess:

$$U_1(m) = \frac{1}{C_1}\left\{\sin k_o ma + \alpha \sin k_1 ma + \left(\frac{\sin k_o a + \alpha \sin k_1 a}{\cos k_o a - \cos k_1 a}\right)(\cos k_o ma - \cos k_1 ma)\right\}$$

as the first solution, and $U_2(m)$ = *same, with* k_o and k_1 permuted. Note that sin(kma) vanishes at m = 0 and so does $\cos(k_o ma) - \cos(k_1 ma)$. The coefficients have, furthermore, been chosen to satisfy the second b.c. U = 0 at m = -1. It follows that U_2 also satisfies both b.c.'s It remains to obtain the normalization constants

$$C_1 = \left[\left(\frac{N_z}{2}\right)\left\{1 + \alpha^2 + 2\left(\frac{\sin k_o a + \alpha \sin k_1 a}{\cos k_o a - \cos k_1 a}\right)^2\right\}\right]^{\frac{1}{2}}$$

with C_2 obtained by the permutation of k_o and k_1. Finally, the parameter α *is chosen to make* U_1 *orthogonal to* U_2. *We obtain:*

$$\alpha = -\left[2 \sin k_o a \sin k_1 a\right]^{-1} x$$

$$\left[\sin^2 k_o a + \sin^2 k_1 a + (\cos k_o a - \cos k_1 a)^2 + \left\{\left[\sin k_o a - \sin k_1 a\right]^2 + \left[\cos k_o s - \cos k_1 a\right]^2\right\}^{\frac{1}{2}} x$$

$$\left\{\left[\sin k_o a + \sin k_1 a\right]^2 + \left[\cos k_o a - \cos k_1 a\right]^2\right\}^{\frac{1}{2}}\Big]$$

Finally, we turn to:

Case c:

Treated as (a) in the range $\epsilon_o \leq E \leq \epsilon_\pi$. From ϵ_π to ϵ_m we follow (b).

It is now important to examine the solutions near the energy minima and maxima. We shall illustrate only the minima: the maxima are treated analogously, so their analysis would be redundant.

We consider the plane index m to be finite, and allow the energy to approach the minimum in case (a) to obtain:

(a) $$U \to i(2/N_z)^{\frac{1}{2}} \sin kma \quad (\text{as } k \to 0)$$

It is important to note that $|U|^2$ vanishes proportional to k^2 as $k \to 0$ for any finite m, but fails to do so if the limit $m \to \infty$ is taken first. Conclusion: square of wavefunction amplitude near energy minimum at surface vanishes, due to b.c. This effect disappears deep within the bulk.

The case (b) is more complex. Defining $\delta \equiv \frac{1}{2}(k_1 - k_o)a$ and proceeding to the limit $\delta \to 0$, we obtain:

(b) $$U \to 2N_z^{-\frac{1}{2}} \cos k_m a \sin k_m am \sin \delta m \quad (\delta \to 0)$$

in which $k_m a = \cos^{-1}(-K/4L)$ (cf. Figure b). U_1 and U_2 cannot be distinguished at any finite value of the index m (although presumably one of the two solutions has an extra node at $m \doteq \frac{1}{2}N_z$ as compared with the other, to ensure orthogonality). Note that, once again, as we approach the energy minimum, $|U|^2 \to 0$, now proportional to δ^2, at finite m. Similar results obtained at the minimum of (c), and at the maxima in all cases. We conclude that the vanishing of the square of the wavefunction amplitudes near the band extrema at the surface of a solid is a general feature of this band structure (first- and second-nearest neighbor overlap).

A Conjecture:

We believe that the above conclusion is a general feature of arbitrary band structures, required because of some as-yet-undiscovered

theorem. A heuristic proof may be given as follows: an arbitrary band structure having a minimum at k = 0 may be approximated by case (a) near the minimum, and the analysis above is then applicable. An arbitrary band structure with a minimum at $k \neq 0$ may be approximated by (b) near the minimum, and the analysis appropriate to that case then applies. In both instances, the square of the amplitudes $|U|^2 \to 0$ proprotional to $(E-\epsilon_{min})$. Computations carried out on several different examples of band structures have confirmed these heuristic arguments.

2D-Brillouin Zone:

For each (k_x,k_y) there is one k_z at which the energy is a minimum and a second at which it is a maximum. Denoting these by $k_z(k_{||})_{min/max}$, we have for the points $k_{||}$ in the nth band, the functions:

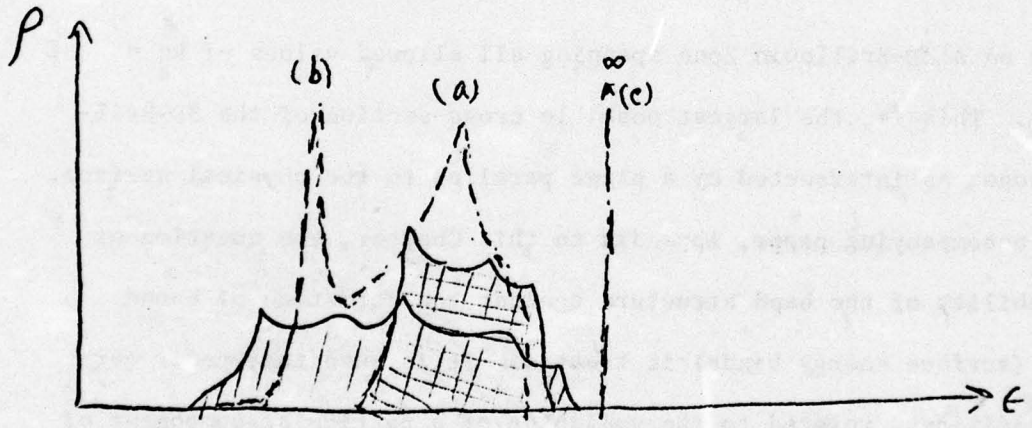
$$\epsilon_n(k_{||},k_z(k_{||})_{min}) \equiv \epsilon_{min}(k_{||}) \text{ and } \epsilon_n(k,k_z(k_{||})_{max}) \equiv \epsilon_{max}(k_{||})$$

defined on a 2D-Brillouin Zone spanning all allowed values of $k_{||} = (k_x,k_y)$. This is, the largest possible cross-section of the 3D-Brillouin Zone, as intersected by a plane parallel to the physical surface. In the accompanying paper, Appendix to this Chapter, the question of the stability of the band structure against the formation of bound states (surface energy bands) is treated. It is seen that under certain conditions, related to the vanishing of a particular component of the bulk inverse-effective-mass tensor, $\alpha_{zz} = 0$, surface states are formed even for infinitesimal surface perturbations. The Brillouin Zone of surface states may cover only a small portion of the maximal 2D-Brillouin Zone. For large surface perturbations, a bound state will form below each energy minimum $\epsilon_{min}(k_{||})$ or for repulsive perturbations,

above each $\epsilon_{max}(k_{\|})$) and the Brillouin Zone of bound states will coincide with the maximal size of a 2D-Brillouin Zone. The reader is referred to the Appendix of this Chapter for examples of the calculational method.

Ad-Atoms on Surface:

We consider the effects of an atomic perturbation on a surface in the presence of surface energy bands. A perturbing potential always has a bound state in 1 or 2 dimensions; however, the bound state here may coincide with the continuum of bulk states, and is broadened into a resonance. This leads to the possibility of the following possible energy levels:



The bulk d.o.s. is shaded, surface states cross-hatched, atomic level dashed line. (a) is broadened atomic level, (b) is bound state below surface band (broadened by resonance with bulk band) and (c) is bound state above both bands, an unbroadened sharp level.

APPENDIX TO CHAPTER III

ELECTRONIC INSTABILITY OF SURFACES OF SOLIDS*

Daniel C. Mattis

Belfer Graduate School of Science
Yeshiva University
New York, N.Y. 10033

Abstract:

We study the effects of a surface perturbation on a semi-infinite solid. The 3D energy band structure is found to determine whether or not there is an intrinsic instability against the formation of surface bands. A criterion, involving one relevant component of the inverse-effective mass tensor, is derived.

*This research supported by a grant N00014-76C-0690 of the Office of Naval Research.

Experimental[1] and theoretical[2] analyses have by now established that the surfaces of solids often differ in crystallographic class[3] as well as in electronic properties[4] from the bulk. The fact that not all surfaces have been found to be reconstructed led me to seek a criterion, and ultimately to pose a simpler problem: under what circumstances will an <u>infinitesimal</u> perturbation produce a band of surface states? We shall denote this an <u>intrinsic</u> instability, for while any arbitrary material may or may not have surface states, those with intrinsic instabilities always <u>must</u>[5]. In the present work, we shall prove that a surface lying in the x-y plane is intrinsically unstable if and only if a component of the bulk inverse-effective mass tensor α_{zz} vanishes at an appropriate point or set of points in the Brillouin Zone. It is thus the bulk band structure and the surface orientation that primarily predetermines the intrinsic surface instability -- the nature of the surface perturbation potential is of secondary importance.

Our prototype material consists of atomic planes at $z = na$, with $n = 1,2,\ldots,N_z$ the plane index. It is terminated by the two surfaces at $n = 0$, N_z+1, with the x and y coordinates continued periodically at N_x+1 and N_y+1 respectively. Proceeding to the limit N_x, N_y and N_z all $\to \infty$, we concentrate on the one surface at $n = 0$. Consider the "unperturbed" Hamiltonian H_o, having the following matrix structure within a single Bloch band:

$$\int d_3r \; \zeta^{+}_{\vec{k}'}(\vec{r}) \; H_o \; \zeta_{\vec{k}}(\vec{r}) = \epsilon_{\vec{k}} \; \delta_{\vec{k},\vec{k}'} \qquad (1)$$

in which ζ_k is the appropriate linear combination of Bloch functions ϕ_k, which satisfies the boundary condition $\zeta_k = 0$ for $z \leq 0$. In the case of a simple band structure having $\epsilon_{\vec{k}}$ (at fixed k_x, k_y) a minimum at $k_z a = 0$ or π, this linear combination is merely:

$$\zeta_k = \frac{1}{\sqrt{2}} \left[\phi_{k_x,k_y,k_z}(r) - \phi_{k_x,k_y,-k_z}(r) \right] \qquad (2)$$

and the first layer amplitudes are $\zeta_k(x,y,a) = \chi_k(x,y) f(k_z a)$ with $f(k_z a) = \sin k_z a$,

In cases when the band structure $\epsilon(k_x,k_y,k_z)$ has a minimum at $k_z a \neq 0$ or π (with $k_\parallel \equiv (k_x,k_y)$ fixed), the amplitude near the surface may be obtained from a solution of the Schrödinger-like equation:

$$\epsilon(k_\parallel, i\frac{\partial}{\partial z})\, f(z) = \epsilon(k_\parallel, k_z)\, f(z), \text{ for } z > 0, \qquad (3)$$

subject to the boundary conditions $f(z \leq 0) = 0$. We have established[6] that for k_z near the energy minimum $f \propto \sin(k_z - k_{z\,min})z$, and obtained a similar result, $f \propto \sin(k_z - k_{zmax})z$, near the energy maximum. These are basically the only properties we shall require.

We next assume that, either as a result of a nonzero change in Madelung potential[5] at small values of z, or because of a small displacement of the surface atoms from their ideal positions, or a result of any other physical requirement, a perturbation Hamiltonian H_s exists near the surface. Explicitly factoring out the amplitudes $f(k_z a)$ of the scattered waves, as dictated by geometric considerations, we may without loss of generality write the matrix structure of H_s in the form:

$$\int d_3 r\ \zeta^{\dagger}_{\vec{k}'}(\vec{r})\ H_s(\vec{r})\ \zeta_{\vec{k}}(\vec{r}) = N_z^{-1}\ g_{k,k'}\ \delta_{k_\parallel,k'_\parallel}\ f(k'_z a)\ f(k_z a) \qquad (4)$$

diagonal in $k_\parallel$ because of translational invariance in x and y. If the perturbation is restricted to the first surface layer (z=a) the coupling constant is a function only of $k_\parallel$, and is denoted $g(k_\parallel)$. If the perturbation affects several surface layers the coupling constant will depend

also somewhat on k_z although for simplicity in the ensuing calculations we approximate it by its value at k_{zmin} (for attractive perturbations) or k_{zmax} (repulsive) so that $g_{k,k'}$ is then just a function of $k_\parallel$, again denoted $g(k_\parallel)$. In fact our criterion for inherent instability does not depend on the value of g, merely on its sign and on the fact that at $k_{zmin/max}$ the coupling constant is neither identically zero nor infinite. Ultimately, if our procedure is found to lead to interesting results, it may be rendered quantitative by obtaining $g_{k,k'}$ self-consistently[8].

We first study the bound states of the joint Hamiltonian $H_o + H_s$, which if they exist, constitute the much discussed[1,4] "surface energy bands". They must be of the form

$$\psi_{k_\parallel,s} = N_z^{-\frac{1}{2}} \sum_{k_z} F(k_\parallel,k_z)\ \zeta_{k_\parallel,k_z} \tag{5}$$

in which we use "s" as a subscript to distinguish these surface modes from the continuum of bulk states. Schrödinger's equation $(H_o + H_s)\ \psi_s = E_s\ \psi_s$ yields the following equation for the coefficients F:

$$(\epsilon_{\vec{k}} - E_s)F(k_\parallel,k_z) + \frac{1}{N_z}\sum_{k_z'} gf(k_z a)F(k_\parallel,k_z')f(k_z' a) = 0 \tag{6}$$

In order for this equation to have a nontrivial ($F \neq 0$) solution, the following secular equation must be satisfied:

$$g^{-1}(k_\parallel) = \Sigma_{1,1}(E_s;k_\parallel) \tag{7}$$

in which $\Sigma_{1,1}$ is a special case of a more general function:

on the surface orientation which defines the z-axis relative to the invariant crystal axes. It is convenient to write the bound state energy in the following analogous form:

$$E_s = \epsilon_{k_\parallel, k_{zmin}} + \alpha_{zz}(k_\parallel, k_{zmin})\ (1-\cosh\ \lambda) \qquad (11)$$

Now, Eq. (7) simplifies to:

$$-\frac{1}{g}\,\alpha_{zz}(\vec{k}_\parallel, k_{zmin}) = \frac{1}{\pi}\int_0^\pi d\theta\ \frac{\sin^2(\theta)}{\cosh\ \lambda + \cos\ \theta} = e^{-|\lambda|} \qquad (12)$$

Thus the r.h.s. of Eq. (12) fails to diverge when $\lambda \to 0$ (in fact, approaches a maximum limiting value of 1). There may nevertheless be an incipient instability against the formation of surface states at arbitrarily small $|g|$ <u>provided there exists a locus in 2D $k_\parallel$-space for which</u> $\alpha_{zz} = 0$. Examples of this include the elementary tight-binding s-bands $\epsilon_{\vec{k}} = -K(\cos k_x a \cos k_y a \cos k_z a)$ for b.c.c. and $-K(\cos k_x a \cos k_y a + \cos k_x a \cos k_z a + \cos k_y a \cos k_z a)$ for f.c.c. crystal structure, both having a minimum at $k_z a = 0$ or π. For each there are rectilinear segments, along which $\alpha_{zz}(\vec{k}_\parallel, 0) = 0$. Therefore, (12) will have solutions along a neighborhood of these lines, over a surface, the area and geometry of which will be functions of the magnitude of g. Conversely, for the s.c. tight-binding band structure $-K(\cos k_x a + \cos k_y a + \cos k_z a)$ for which the minimum is also at $k_z = 0$, there exists no curve over which $\alpha_{zz} = 0$, and therefore no region of $\vec{k}_\parallel$-space inherently unstable against the formation of surface states.

The analysis for $g \geq 0$ is similar. The Bloch energy is expanded about its maximum, the effective mass parameter is now non-positive, and the bound state lies above the continuum. With the aid of

$$\Sigma_{n,m}(E;k_{\|}) \equiv N_z^{-1} \sum_{k_z'} \left\{ \frac{f(k_z'na)\ \ f(k_z'ma)}{E - \epsilon_{k_{\|},k_z'} - i0^+} \right\}$$

$$= \frac{a}{\pi} \int_0^{\pi/a} dk_z'\ f(k_z'\,na)\ f(k_z'ma) \left\{ \frac{P.P.}{E-\epsilon_{k_{\|},k_z'}} + i\pi\delta\ (E-\epsilon_{k_{\|},k_z'}) \right\} \quad (8)$$

Normalization of the wave function requires:

$$F(\vec{k}) = \frac{f(k_z a)}{E_s - \epsilon_{\vec{k}}} \left[\frac{-\partial\ \Sigma_{1,1}(E_s;k_{\|})}{\partial E_s} \right]^{-\frac{1}{2}} \quad (9)$$

As E_s approaches $\epsilon(k_{\|},k_{zmin})$ from below (for attractive potentials, or as E_s approaches $\epsilon(k_{\|},k_{zmax})$ for repulsive) the denominators become arbitrarily small. However, the numerators also vanish (cf. discussion following Eq. (3))[6] and thus the integrals remain finite unless — due to the peculiarities of the band structure — the denominator vanishes as $(k_z-k_{zm})^4$. The conditions for this to occur are now examined. For $g < 0$, we start by expanding the Bloch energy about its minimum:

$$\epsilon_{k_{\|},k_z} \simeq \epsilon_{k_{\|},k_{zmin}} + \tfrac{1}{2}\,\alpha_{zz}(k_{\|},k_{zmin})(k_z-k_{zmin})^2\ a^2 + 0((k_z-k_{zmin})^4)$$

or, more compactly, with $\theta \equiv (k_z-k_{zmin})a$,

$$\epsilon_{k_{\|},k_z} \simeq \epsilon_{k_{\|},k_{zmin}} + \alpha_{zz}(k_{\|},k_{zmin})(1-\cos\,\theta) + 0\ (\theta^4) \quad (10)$$

in which $\alpha_{zz}(k_{\|},k_{zmin}) \equiv (m^*)^{-1}_{zz} \equiv a^{-2}\partial^2\epsilon_{\vec{k}}\,/\partial k_z^2$ is the zz-component of the inverse-effective-mass tensor. Because it is evaluated at a minimum, $\alpha_{zz} \geq 0$ necessarily; its magnitude depends both on the band structure and

$$\epsilon_{\vec{k}} \simeq \epsilon_{k_\parallel, k_{zmax}} + \alpha_{zz}(k_\parallel, k_{zmax})(1 + \cos\theta) + 0(\theta^4)$$

and

$$E_z = \epsilon_{k_\parallel, k_{zmax}} + \alpha_{zz}(k_\parallel, k_{zmax})(1-\cosh\lambda), \qquad (13)$$

Eq. (7) once again reduces to (12), after the substitution of $\alpha_{zz}(k_\parallel, k_{zmax})$ for $\alpha_{zz}(k_\parallel, k_{zmin})$ in the l.h.s. Because of the change in sign of the effective-mass parameter, the bound-state solutions now exist only for $g \geq 0$, but aside from this, the discussion given after Eq. (12) applies for this case also. It should be noted that the loci of $\alpha_{zz} = 0$ are not necessarily the same for k_{zmax} as for k_{zmin}, although in the two examples given (b.c.c. and f.c.c.) they do coincide because of symmetry.

The s.c. tight-binding band structure provides an example of the importance of surface orientation. Inherently stable against the formation of surface bands when the surface is a (100) plane, it demonstrates an inherent instability along a (011) direction: in the new coordinate system $a_y = a_z = a2^{-\frac{1}{2}}, a_x = a$, and the Bloch energy takes the form

$$\epsilon_{\vec{k}} = -K(\cos k_x a + 2\cos k_y a_y \cdot \cos k_z a_z)$$

For $k_{zm} a_z = 0, \pi$ the equation $\alpha_{zz} = 0$ has a solution along the straight-line segments k_x = arbitrary, $k_y = \pm \pi/2a_y$, indicating an instability against the formation of surface states in the neighborhood of these lines.

A surface perturbation scatters bulk states, and modifies the surface density of states g affecting optical absorption at the surface, etc., even when there are no states bound to the surface. The contributions of bound- and scattered- states are combined in the calculation of the "perturbed" local density of states function[2] on the n-th plane. After some algebra, one obtains:

$$\rho_n(\omega) = (N_x N_y)^{-1} \sum_{\vec{k}_\parallel} (\Sigma^2_{1,n}(E_s;k_\parallel)) \; (-\partial\Sigma_{1,1}(E_s;k_\parallel)/\partial E_s)^{-1} \delta(\omega - E_s)$$

$$+ 2\,(N_x N_y N_z)^{-1} \sum_{\vec{k}} |\Lambda_{1,n}(\vec{k})|^2 \; \delta(\omega - \epsilon_{\vec{k}}) \qquad (14)$$

where E_s is the surface-band energy, and

$$\Lambda_{1,n}(\vec{k}) \equiv \frac{f(k_z na)\left[(1-g(k_\parallel)\Sigma_{1,1}(\epsilon_{\vec{k}};k_\parallel)\right]+g(k_\parallel)f(k_z a)\Sigma_{1,n}(\epsilon_{\vec{k}};k_\parallel)}{1 - g(k_\parallel)\;\Sigma_{1,1}(\epsilon_{\vec{k}};k_\parallel)} \qquad (15)$$

In the bulk limit $n \to \infty$ we drop the exponentially decaying and/or rapidly oscillating terms to obtain the limiting behavior $\rho_\infty(\omega) = (N_x N_y N_z)^{-1} \sum_{\vec{k}} \delta(\omega - \epsilon_{\vec{k}})$, which is exactly the usual bulk value. Thus, the existence of the surface is "forgotten" deep within the bulk. For small n however, the local density-of-states given by (14) is sensitive to the sign and strength of the perturbing potential, and to the existence of bound states.[10]

The modified charge distribution is conducive to the phenomenon of "surface reconstruction"[3,8]. We have seen that the surface states (if any) do not necessarily occupy the full 2D surface Brillouin Zone, but just a small area near the locus of the instability $\alpha_{zz} = 0$. The narrowness of this area in $\vec{k}$-space suggests a large resultant cell size in real space, as is indeed experimentally observed in some reconstructed surfaces. The predicted absence of inherent surface-state instability in other cases (e.g. the model s.c. with (100) surface) suggests the absence of surface reconstruction there. In all cases, an analysis of $\alpha_{zz}(k_\parallel, k_{zm})$ to determine where (or if) $\alpha_{zz} = 0$ might usefully precede any full-blown study of a surface, to pin-point the a-priori instabilities.

I thank Drs. Rafael Pena and Barry Simon for helpful collaboration.

Footnotes and References

1. See the various articles on surface physics in Physics Today, vol. 28, No. 4, April 1975. Also: J.A. Appelbaum and D.R. Hamann, Revs. Mod. Phys. 48, 479 (1976). Also, "Surface Physics of Phosphors and Semiconductors", C.G. Scott and C.E. Reed, Editors, Academic Press, N.Y., 1975.

2. The theoretical analyses and concepts (e.g. "local density-of-states") are reviewed in J.R. Schrieffer and P. Soven, Physics Today 28, 24 (1975). (Note that they number the first atomic plane n=0, whereas in the present work it is n=1). See also Chapters by R.O. Jones and by F. Berg in Scott and Reed, op. cit.

3. P.J. Estrup, Physics Today 28, 33 (1975).

4. such as, the formation of bands of energy levels confined to the neighborhood of the surface; see J.A. Appelbaum and D.R. Hamann, Phys. Rev. Letters 32, 225 (1974), as well as Eqs. (4)-(11) in the present paper.

5. Because surface perbutations always exist. For example, the ubiquitous Madelung potential necessarily deviates from its bulk value in the neighborhood of a surface, as emphasized by A. Clark in his "The Chemisorptive Bond", Academic Press, New York, 1974 §9.3.

6. We have completely solved two examples: $\epsilon = (\frac{\partial^2}{\partial z^2} + k^2_{zmin})^2$, with the helpful collaboration of Prof. Barry Simon, and $\epsilon = -A \cos a\frac{\partial}{\partial z} + B \cos 2a \frac{\partial}{\partial z}$. In both instances f(z) has the behavior indicated in the text, and we have compelling reason to believe that its vanishing near k_{zmin} (or k_{zmax}) is a very general property; a full account will be published elsewhere.

7. Mathematically, the problem of obtaining bound- and scattering-states of $H_o + H_s$ remains exactly soluble for any $g_{k,k'}$ which, while an arbitrary function of $k_{||}$, is of the separable form $\phi(k_z)\phi(k_z')$ in k_z, or is a finite sum of each terms. Instead of (7) we obtain a secular determinant containing the quantities $\Sigma_{n,m}$. The condition for the existence of a bound state when $|g| \to 0$ remains precisely $\alpha_{zz} = 0$ in this case. In the most general case, $g_{k,k'}$ an arbitrary function of k_z and k_z', we once again recover the same criterion using a variational solution for the bound state. This requires only that $g_{k,k'}$ be suitably bounded — neither zero nor infinite at k_{zm}.

8. As, for example, in the work of M. Schlüter, J.R. Chelikowsky, S.G. Louie and M.L. Cohen, Phys. Rev. Letters 34, 1385 (1975). In our method, the calculation of a self-consistent surface perturbing potential H_s would involve the perturbed density-of-states Eq. (14) and Poisson's equation.

9. as observed in photoelectronic spectroscopy: see D.E. Eastman and M.I. Nathan, Physics Today 28, 44 (1975) and R.L. Park, ibid., p. 52.

10. The reader should note that the continuum of surface states in general overlaps the bulk band continuum of Bloch energies. Thus the surface bands, if any, may lie entirely within the Bloch-state band (this is the case of an intrinsically unstable surface under the influence of a very weak surface perturbation), or may only partly overlap it (somewhat stronger perturbing potential) or may indeed lie entirely outside the bulk band (as happens when the perturbing potential is large, regardless of whether or not the surface is intrinsically unstable in the sense of the present work.)

IV. TRANSPORT IN A RANDOM MEDIUM

Impurity-band conduction and electron energy bands in random alloys may be studied by similar techniques. The best single method of the past decade has been the CPA, which suffers from a basic weakness: inability to take short range order (SRO) into account. A relatively simple method, yet one that can account for SRO, and LRO if necessary, is outined in the Appendix to this Chapter. Extension of this to the Hubbard model (2-body forces) is actively underway, the SRO being determined self-consistently.

APPENDIX TO CHAPTER IV

ELECTRON STATES IN RANDOM ALLOYS WITH SHORT-RANGE ORDER

Paul Bloom* and Daniel Mattis#

Belfer Graduate School of Science

Yeshiva University,
New York, N.Y. 10033

Abstract:

We present an accurate and economical iterative method of calculating the energy levels of a disordered or partly ordered random alloy. Results presented for 1-D and 3-D simple-cubic lattices compare favorably with exact calculations. We also present the systematic effects of partial short-range order in 3-D. A theory of the one-particle propagators is presented, and the theory of electrical conductivity is developed in the context of our new method. Our formulas satisfy the exact conservation laws.

* Supported by AFOSR Grant #73-2430B.

This research is supported by a grant from the Office of Naval Research #N00014-76-C-0690

I Introduction

The study of electronic and vibrational spectra of disordered alloys is currently one of the principal concerns of solid state physics[1], stimulated by the outstanding successes of the coherent potential approximation[2,3] (CPA), now ending its first decade. Although efforts to improve our understanding beyond the CPA have not all met with the same good fortune, there have been recent exceptions. Cluster methods[4,5] have been devised which are accurate enough to reproduce the "peaky" structure of the density of-states $\rho(\omega)$, which they sometimes do (notably in one dimension[4]) with startling fidelity. We have been working along such a cluster-type approach, and have found an extremely simple method translating directly into a computer algorithm. While unsuited to the theoretical study of Lifshitz[6] tails, our method has permitted us to reproduce many of the other known results over the theoretically permitted range of energy[7], even near the energy maxima and minima, and additionally, permitted to study of the effect of short-range order (SRO). Along with Lifshitz, we envisage tails in $\rho(\omega)$ at the energy maxima and minima as arising from accidental correlations in increasingly large clusters, of a size that for practical reasons we are not at present capable of handling; however, the simplicity of the present method may suggest a natural extension to cover this[8].

The basic outline of our paper is as follows: In Section II we present a method for the calculation of the single-body Green function in the presence of an arbitrary number of impurities. We then discuss how our procedure can be implemented by the use of a convergence factor, Σ. Section III is devoted to an analysis of the

meaning and uses of the complex self-energy Σ within the context of a disordered medium. Results from our method are presented in Section IV, including the effects of SRO. Beyond this in Section V we make further approximations that allow us to determine $\overline{G_{\vec{k}\,\vec{k}}}(\omega)$. Section VI is concerned with the development of a transport theory compatible with $\bar{G}$, along the lines of Baym and Kadanoff.[9]

II Cluster Green Function

Let the Hamiltonian for the electrons within a single tight binding band in a hypercubic lattice in D dimensions be:

$$H = \sum_{i,j} T_{i,j} |i\rangle\langle j| + \sum_{i} V_i |i\rangle\langle i| \equiv T + V \tag{1}$$

with $T_{ij} = (2D)^{-1}$ for i,j nearest-neighbors and zero otherwise, $|i\rangle$ the Wannier state at the lattice point R_i, and V_i the potential which takes on one of two values depending whether atom A or B occupies the i site. We construct the resolvent operator G(z) and its various matrix elements:

$$G(z) \equiv (z-H)^{-1} = [z - (T+\Sigma) - (V - \Sigma)]^{-1} \tag{2}$$

in which we reference the operators to a complex "optical potential" $\Sigma(z)$ merely as a device to enhance the convergence of subsequent expansions, with z = frequency ω, extended to the complex plane.

For those readers familiar with the CPA, it is important to note that our new departure consists principally in dissociating the complex self-energy parameter $\Sigma(\omega)$ from the site-diagonal averaged Green function $\overline{G_{nn}}(\omega)$. Whereas in CPA, knowledge of the one implies the other, via the relationship:

$$\overline{G_{nn}}_{CPA}(z) = \langle n|[z - (T + \Sigma_{CPA}(z))]^{-1}|n\rangle \ , \tag{3}$$

our experience indicates that it is better to treat $\Sigma(\omega)$ merely as a convergence parameter, one to be chosen as an ad-hoc aid in the calculations rather than by tedious and unnecessary self-consistency conditions. As by Eq. (2) the exact $\overline{G_{nm}}(z)$ are all independent of $\Sigma(z)$, in any accurate approximation to $\overline{G_{nm}}(z)$ we have latitude in our choice of $\Sigma(z)$, as discussed below, and we pick the simplest possible $\Sigma(z)$ for which our calculated G is approximately stationary.

We next define a modified resolvent operator, $G^{(i)}$, appropriate to the case in which one sets $\tilde{V}_i = 0$, where we define $\tilde{V}_i \equiv \langle i|(V-\Sigma)_i|i\rangle$ and, indicating the elimination of the localized fluctuation potential at this site by $(\quad)'_i$, we have:

$$G^{(i)}(z) \equiv [z - (T + \Sigma) - (V - \Sigma)'_i]^{-1} \ . \tag{4}$$

The full resolvent (2) can be expressed in terms of the modification in (4) by the use of the operator identity $(A-B)^{-1} = A^{-1} + A^{-1}\,B(A-B)^{-1}$;

$$G(z) = G^{(i)}(z) + G^{(i)}(z)\,(V - \Sigma)_i\,G(z). \tag{5}$$

Because the perturbation is diagonal in the Wannier representation, the matrix elements are easily found:

$$G_{nm}(z) = G^{(i)}_{nm}(z) + \frac{G^{(i)}_{ni}(z)\,\tilde{V}_i\,G^{(i)}_{im}(z)}{1 - G^{(i)}_{ii}(z)\,\tilde{V}_i} \quad . \tag{6}$$

For the calculation of the density of states function $\rho(\omega) \equiv \frac{-1}{\pi}\,\mathrm{Im}\overline{G_{nn}}(\omega+i\varepsilon)$ only the configurationally-averaged $\overline{G_{nn}}(\omega+i\varepsilon)$ is required. For the one-particle propagators $\overline{G_{\vec{k}\vec{k}}}$ the averaged Fourier transforms of all $\overline{G_{nm}}\,(\omega + i\varepsilon)$ are needed. Eq. (6) is now iterated. Define $G^{(i,j)}(z)$ to be the modified resolvent operators with the fluctuation-potentials $\tilde{V}$ at sites i and j removed. By a repetition of the above, we have:

$$G^{(i)}_{nm}(z) = G^{(i,j)}_{nm}(z) + \frac{G^{(i,j)}_{nj}\,\tilde{V}_j\,G^{(i,j)}_{jm}}{1 - G^{(i,j)}_{jj}\,\tilde{V}_j} \quad . \tag{7}$$

The matrix elements G_{nm} decay exponentially with distance R_{nm}; thus the expansion (6), (7) is in a symbolic "parameter" γ defined as $G_{nj}{}^{(\ldots)}\tilde{V}_j$, which is "small" for small $\tilde{V}_j$ and "exponentially small" at large $\tilde{V}_j$. The process (6), (7) is to be repeated any numer of times, until the largest practical cluster size is achieved.[10] Termination, by truncation, of the series consists of approximating the most distant G's, i.e. those with the largest number of superscripts, by their value in the average optical potential. Thus, if we stop at (7), the approximation consists in replacing $G^{(i,j)}_{nm}(z)$ by $\langle n|[z - (T + \Sigma)]^{-1}|m\rangle$. The configurational averages over all the explicitly retained V_i are then performed, and all $\bar{G}$'s obtained.

III Choice of Σ

We now come to our principal point of departure from other methods - our choice of Σ. Our results would depend crucially upon Σ except for the following observations. Since the behavior of the local cluster is the dominant characteristic of disordered systems, we expect results insensitive to the particular choice of Σ if the cluster size is sufficiently large.

We require a simple functional form for Σ that allows for states out to the bands limits. This excludes the use of Σ_{CPA}, which is known to produce bands that are always too narrow. We restrict the range of possible Σ's by requiring that it obey dispersion relations, insuring that our approximate G is analytic. Furthermore, a functional form is desired in which G is accurate in both the weak as well as strong scattering regimes. Because of the local nature of highly disordered systems, our choice becomes more critical for small potential differences where effects are more extended. Our input is the Im Σ which we take as one or more step functions, non-zero only within the theoretical band limits. Re Σ is then determined from the following dispersion relation

$$\Sigma(z) = \frac{\langle V \rangle}{2} + \frac{1}{2\pi i} \int_{-\infty}^{\infty} dx \, \frac{\Sigma(x+i\epsilon)}{x-z} \quad . \tag{8}$$

This is sufficient to make our approximation to G(z) satisfy causality. The density-of-states sum rule, $\int_{-\infty}^{\infty} d\omega \, \rho(\omega) = 1$ is itself a beneficial consequence of the analyticity of our approximate G(z) and its resulting 1/z dependence in the asymptotic limit as we discuss elsewhere.[11] We verified that in all cases studied, the sum rule on $\rho(\omega)$ was satisfied numerically.

If we choose the constant, Im Σ, to be of magnitude of Im Σ_{CPA} then in the weak scattering and low concentration regimes G(z) will be quite similar to $G_{CPA}(z)$ so that accurate results can be expected in all regimes.

Before we proceed, the way in which we use Im Σ_{CPA} must be more clearly outlined. In the accompanying Figure 1 we display the two basic behavior patterns of $|\text{Im } \Sigma_{CPA}|$ as observed by VKE.[12] It should be noted that, here also, ReΣ and ImΣ are related by the Eq. (8). Σ has to describe evervthing in the CPA; it determines band gaps, peaks in the density-of-states, and the general overall scale. Most of these results (e.g. band gaps and complicated structure) are better obtained by our detailed calculations of the correlated scattering. We hypothesize that the most useful information from CPA is contained in the general overall magnitude of Im Σ_{CPA}. Operationally, in Figure 1a, we would ignore values of $|\text{Im } \Sigma_{CPA}|$ from the region of its maximum as well as the extremeties of the band. In the former range of energies, we expect exceptional scattering because it is easiest for these states to make transitions due to band overlap, whereas at the band edges the spectrum will be least disturbed, according to the same considerations. Any value from the shaded region is then acceptable. In terms of particle lifetimes, we will obtain the large and small transition rates because we almost solve the eigenvalue problem exactly for each configuration and this is clearly equivalent to a perturbation approach. As for concentration dependence in Σ_{CPA}, we will obtain correct behavior simply because we weigh each configuration by its appropriate probability. Thus we are able to include both the dynamical and statistical aspects of the problem.

In case b, the same analysis leads us to ignore the very large values of $|\text{Im } \Sigma_{CPA}|$ in both subbands. Here though, the magnitudes are considerably different leading us to suspect that two different constants are needed. Further details of this case will be elucidated in the following examples.

IV Analysis of Results

We first consider the canonical 1-D tight binding binary alloy for 3 different scattering strengths at a 50-50 concentration. Figure 2 compares the results of 1,3, and 5 cluster calculations for $\rho(\omega)$ when $V_i = \pm .5$ with exact results. We see in this example the development of the peaky structure associated with special clusters of atoms as our cluster size increases. Proceeding to a larger scattering strength ($V_i = \pm 1.0$), we expect that the local configurations will play a more prominent role because of increased wave function localization. As shown in Figure 3 we successfully reproduce most of the structural details of $\rho(\omega)$ for a 5 cluster. To check the degree of insensitivity in our 5 cluster model we varied $|\text{Im } \Sigma|$ within the limits given by $\text{Im } \Sigma_{CPA}$ and found little change in the overall pattern as shown in Figure 4 This indicates, numerically that the resulting G is stationary and that Σ is optimum.

The scattering strengths are now increased to $V_i = \pm 2.0$, providing a critical evaluation of the methods capabilities (larger scattering strengths are in a sense too easy because wave function localization makes a cluster calculation more plausible). Using the exact scattering off all configurations of 5 atoms, the highly discrete

spectrum is well reproduced, as seen in Figure 5a. Increasing the cluster size to 7 atoms, keeping the convergence factor the same as in 5A, improves our agreement with the exact results as shown in Figure 5b. In Figure 6 we display the results of again varying $|\text{Im } \Sigma|$ within the limits dictated by Im Σ_{CPA}; the major details are again seen to remain stationary. We have found empirically that if $|\text{Im } \Sigma|$ is too small, the resultant density-of-states is too "peaky" and as such, representative of a molecular cluster, instead of the solid state. If $|\text{Im } \Sigma|$ is too large, then the central site predominates, as is correct only in the extreme "atomic" limit when potential fluctuations greatly exceed the band-width. One can see this from Figure 6 since the sharper curve is associated with the lowest value of $|\text{Im } \Sigma|$ and vice-versa.

In three dimensions the obvious cluster size is 7 sites. Figure 7 compares our calculation with the Monte-Carlo type numerical results of Alben et. al.[13] A constant Im Σ gave poor results in this case, but the CPA calculations immediately showed us why : Im Σ_{CPA} was more than one order-of-magnitude smaller in the majority subband than in the minority subband. Consequently we changed Im Σ to the step function shown in the figure, varying the parameters (magnitudes of the steps) again guided by CPA. The results now agreed well with the exact computations and were insensitive to the precise value of our parameters as is evidenced by Figure 8 in which a 3-step function was used.

To illustrate entirely new applications, consider effects of SRO on this same alloy. With α the Cowley SRO parameter, c_A and $c_B = 1 - c_A$ the relative concentrations and P_{AB} the probability of finding atom A at a given site when a B atom occupies a specified neighboring site, we have $P_{AA} = c_A + c_B \alpha$, $P_{BA} = c_B(1 - \alpha)$, $P_{AB} = c_A(1 - \alpha)$, and $P_{BB} = c_B + c_A\alpha$. In ref. 13, α=0. For $c_A = .1$, α can vary from -0.11 to +1.0; negative α is associated with enhanced tendency of A atoms to be surrounded by B's (i.e. "antiferromagnetism"), positive α indicates enhancement in the probability of either species being surrounded by atoms of its own kind (i.e. "ferromagnetism"). Using the same convergence parameters as in our calculation at $\alpha = 0$, In Figure 9 we find distinctive features in the minority sub-band density of states that we interpret in terms of minority-atom clustering: the single peak of $\alpha = -.07$ registers the unlikelyhood of finding two A atoms as nearest neighbors, and the double peaks of $\alpha = .7$ represent the tendency of the same atoms to form pairs, triplets, etc. However, due to the sparseness of A atoms, triplets and higher-order clusters are statistically insignificant for these values of α.

V Electron Propagation

So far we have developed a method for calculating the site-diagonal configuration averaged Green function. We have not indicated, how we would calculate the non site-diagonal propagators. One alternative is to develop a cluster method for the latter, similar to the method we used for the former. Another, simpler though less accurate, alternative will be employed. We first define a new self-energy $\Lambda^*(z)$ by the equation

$$\overline{G_{\ell\ell}}(z) \equiv \frac{1}{N} \sum_{\vec{k}} \frac{1}{z - \Lambda^*(z) - \epsilon_{\vec{k}}} . \tag{9}$$

This relationship is numerically inverted to obtain Λ^* as a function of the exact or numerically calculated $\overline{G_{\ell\ell}}$. From the calculation of $\Lambda^*(\omega)$ we obtain $\Lambda^*(z)$ in the entire complex plane, as:

$$\Lambda^*(z) = \langle V \rangle + \frac{1}{\pi} \int_{-\infty}^{\infty} d\omega \frac{\mathrm{Im}\, \Lambda^*(\omega+i\epsilon)}{\omega - z}. \tag{10}$$

If further values of $\overline{G_{\ell m}}(z)$ were calculated numerically, then we would determine $\Lambda^*(\vec{k},z)$ from

$$\overline{G_{\ell m}}(z) \equiv \frac{1}{N} \sum_{\vec{k}} \frac{e^{i\vec{k}\cdot(\vec{R}_\ell - \vec{R}_m)}}{z - \Lambda^*(\vec{k},z) - \epsilon_{\vec{k}}} ,$$

and (11)

$$\overline{G_{\vec{k}\vec{k}}} = \frac{1}{z - \Lambda^*(k,z) - \epsilon_{\vec{k}}} .$$

The analysis is facilitated by going over to a localized representation in which we would specify the number of elements $\Lambda^*_{\ell m}(z)$ that we have determined numerically. For example, if we have available $\overline{G_{\ell,\ell}}(z)$, $\overline{G_{\ell,\ell+1}}(z)$, and $\overline{G_{\ell,\ell+2}}(z)$, then we would be able to obtain

$$\Lambda^*_{\ell,m}(z) = \delta_{\ell,m} \Lambda^*_0(z) + \delta_{\ell+1,m} \Lambda^*_1(z) + \delta_{\ell+2,m} \Lambda^*_2(z),$$

by solving the 3 equations simultaneously. In the case at hand, we will use a site-diagonal self-energy since all we have at our disposal are the computed $\overline{G_{\ell\ell}}(z)$. We will still use

$$\overline{G_{\ell m}}(z) = \frac{1}{N} \sum_{\vec{k}} \frac{e^{i\vec{k}\cdot(\vec{R}_\ell - \vec{R}_m)}}{z - \Lambda^*(z) - \epsilon_{\vec{k}}} ,$$

hence

$$\overline{G_{\vec{k}\vec{k}}} = \frac{1}{z - \Lambda^*(z) - \epsilon_{\vec{k}}} \qquad (12)$$

as the definition of the off-diagonal elements. The propagators decay rapidly with distance $R_{\ell m}$ so both the one, two, or three point curve fitting proceedures will probably give reasonably equivalent results. Now, all the information contained in our previous numerical work is stored in $\Lambda^*(z)$, the complex proper self-energy part. It is of interest to compare Im Λ^* with Im Σ_{CPA} in order to see how they differ. This is done in Figure 10 for the 1-D alloy of Figures 5 and 6.

In summary we have presented a relatively simple method for calculating the eigenvalue spectrum of a disordered system, one that avoids all the computational pitfalls of self-consistent methods. This quasi-invariant theory is not only highly accurate, but also allows the bounds on the frequency spectrum to be naturally determined by the correlated scattering of a local group . We now discuss transport and develop a formalism that allows our numerical output to be used in approximations that conserve particle number and energy.

VI Transport in Disordered Systems

The linear response of the current to the electric field defines the conductivity, which we take to be the same along the 3 principal directions in our simple cubic structure. Following Velický[14] we have in our single-particle model

$$\sigma(0) = \frac{\pi}{m^2} \int_{-\infty}^{\infty} d\lambda \quad \left[-\frac{\partial f(\lambda)}{\partial \lambda}\right] \quad \overline{\langle \delta(\lambda-H) p_1 \delta(\lambda-H) p_1 \rangle} \tag{13}$$

where e , the electric charge, is unity and f is the fermi function. The bracketed term is short hand for

$$\overline{\langle \delta(\lambda-H) p_1 \delta(\lambda-H) p_1 \rangle} \equiv \overline{\sum_{\alpha} \langle \alpha | \delta(\lambda-H)\, p_1\, \delta(\lambda-H)\, p_1 | \alpha \rangle} \tag{14}$$

p_i is the momentum operator along an arbitrarily chosen principal axis and the long bar denotes configuration averaging. Examination of Eq (14) reveals that we require the two-particle correlation function

$$\overline{G^2_{imj\ell}(z_1,z_2)} = \overline{\langle i | \frac{1}{z_1 - H} | j \rangle \langle m | \frac{1}{z_2 - H} | \ell \rangle} \quad . \tag{15}$$

We can relate $\overline{G^2}$ to $\overline{G}$ by the equation

$$\overline{G^2_{imj\ell}(z_1,z_2)} = \overline{G_{ij}(z_1)}\,\overline{G_{m\ell}(z_2)} + \sum_{\substack{kp \\ qr}} \overline{G_{ik}(z_1)}\,\overline{G_{q\ell}(z_2)}\, \Xi_{kr,pq}\, \overline{G^2_{pmjr}(z_1,z_2)}, \tag{16}$$

which defines the index structure of the vertex function. This equation is exact for the exact $\bar{G}$ and we will use it to define $\overline{G^2}$ when we have an approximate single particle green function. We can place restrictions on possible vertex functions by requiring the conservation of charge and energy in the presence of a long-wavelength disturbance. This leads to the introduction of a new operator

$$K(z_1,z_2) = \overline{\frac{1}{z_1-H}\ \frac{1}{z_2-H}} \quad . \tag{17}$$

It is easy to show that K must satisfy the following Ward-type identity:

$$K(z_1,z_2) = \frac{1}{z_2-z_1}\left[\bar{G}(z_1) - \bar{G}(z_2)\right] \tag{18}$$

which implies a connection between the one-body operator $\bar{G}$ and the two-body green function $\overline{G^2}$. Also, one can show with the above condition that the linear response of the particle number and energy to a long wavelength disturbance is zero, thus ensuring the appropriate conservation laws. If we use an approximate $\bar{G}$, then we must construct a K that maintains Eq. (18) and this allows us to relate $\bar{G}$ to the vertex function in the following way: We let $z_2 \to z_1$, then Eq (18) becomes

$$K_{i\ell}(z_1,z_2) = -\frac{d\,\overline{G_{i\ell}}(z_1)}{dz_1} \,. \tag{19}$$

The configuration averaged resolvent can be written as

$$\overline{G_{i\ell}}(z_1) = \langle i \left| \frac{1}{z_1 - T_{kin} - \Lambda^*_{op}(z_1)} \right| \ell \rangle \tag{20}$$

which leads to the equation

$$K_{i\ell}(z_1,z_1) = \sum_j \overline{G_{ij}}(z_1)\,\overline{G_{j\ell}}(z_1) + \sum_{kq} \overline{G_{ik}}(z_1)\left[\frac{-d\Lambda^*_{kq}(z_1)}{dz_1}\right]\overline{G_{q\ell}}(z_1). \qquad (21)$$

But

$$-\frac{d\Lambda^*_{kq}(z_1)}{dz_1} = \sum_{rp} \frac{d\Lambda^*_{kq}(z_1)}{d\overline{G_{pr}}(z_1)}\left[\frac{-d\overline{G_{pr}}(z_1)}{dz_1}\right] = \sum_{rp} \frac{d\Lambda^*_{kq}(z_1)}{d\overline{G_{pr}}(z_1)} K_{pr}(z_1,z_1)$$

so that

$$K_{i\ell}(z_1,z_1) = \sum_j \overline{G_{ij}}(z_1)\overline{G_{j\ell}}(z_1) + \sum_{\substack{kq \\ rp}} \overline{G_{ik}}(z_1)\overline{G_{q\ell}}(z_1)\,\frac{d\Lambda^*_{kq}(z_1)}{d\overline{G_{pr}}(z_1)}\,K_{pr}(z_1,z_1). \qquad (22)$$

In $\overline{G^2}$, we let m = j and sum over all j with $z_1 = z_2$:

$$\sum_j \overline{G^2_{ijj\ell}}(z_1,z_1) = \sum_j \overline{G_{ij}}(z_1)\overline{G_{j\ell}}(z_1) + \sum_{\substack{kq \\ rp}} \overline{G_{ik}}(z_1)\overline{G_{q\ell}}(z_1)\,\Xi_{kr,pq}(z_1,z_1)\sum_j \overline{G^2_{pjjr}}(z_1,z_1). \qquad (23)$$

Once we recognize $K_{i\ell}(z_1,z_1) = \sum_j \overline{G^2_{ijj\ell}}(z_1,z_1)$, we find that the vertex must satisfy the equation

$$\Xi_{kr,pq}(z_1,z_1) = \frac{d\Lambda^*_{kq}(z_1)}{d\overline{G_{pr}}(z_1)}\,. \qquad (24)$$

We note that not only must this relationship hold for the exact vertex function and self-energy but also in any approximation in

in which it is desired that the two-particle correlation function satisfy the Ward Identity, Eq.(18). This means that once an approximation is made to Λ^*, we can determine transport functions that allow the conservation laws to be obeyed. We could make further approximations to $\overline{G^2}$ but we would then have no guarantee of conserving charge, energy, etc. The Ward Identity is useful to generate a vertex function only when the frequencies are the same. For the case at hand, $\Lambda^*_{kq}(z_1) = \delta_{kq}\,\Lambda^*(z_1)$ is a function of the site-diagonal averaged Green function so the vertex is

$$\Xi_{kr,pq}(z_1,z_1) = \delta_{kq}\delta_{rp}\frac{d\Lambda^*(z_1)}{d\overline{G}(z_1)} = \delta_{kq}\delta_{rp}\,\Xi(z_1,z_1). \tag{25}$$

For $z_1 \neq z_2$, we make the approximation

$$\Xi_{kr,pq}(z_1,z_2) = \delta_{kq}\,\delta_{rp}\,\Xi(z_1,z_2). \tag{26}$$

This is certainly consistent with the Ward Identity, and furthermore, it allows us to show that contributions to the conductivity from the vertex correlations then vanish. The two-particle correlation function is now

$$\overline{G^2_{imj\ell}}(z_1,z_2) = \overline{G_{ij}}(z_1)\overline{G_{m\ell}}(z_2) + \sum_k \overline{G_{ik}}(z_1)\overline{G_{k\ell}}(z_2)\,\Xi(z_1,z_2)\sum_r \overline{G^2_{rmjr}}(z_1,z_2)\ . \tag{27}$$

To find the conductivity we need, $\langle\delta(\lambda_1-H)p_1\delta(\lambda_2-H)p_2\rangle$, or

$$I_{12}(\lambda_1,\lambda_2) = \sum_{ij\ell m} \vec{p}^{\,1}_{jm}\overline{\langle m|\delta(\lambda_2-H)|\ell\rangle}\ \vec{p}^{\,2}_{\ell i}\overline{\langle i|\delta(\lambda_1-H)|j\rangle} \tag{28}$$

which requires

$$\sum_{ij\ell m} p^{\vec{1}}_{jm}\, p^{\vec{2}}_{\ell i}\, \overline{G^{2}_{imj\ell}}(\lambda_1,\lambda_2) = \sum_{ij\ell m} p^{\vec{1}}_{jm}\, p^{\vec{2}}_{\ell i}\, \overline{G_{ij}}(\lambda_1)\, \overline{G_{m\ell}}(\lambda_2)$$

$$+ \sum_{ij\ell m}\sum_{k} p^{\vec{2}}_{\ell i}\, \overline{G_{ik}}(\lambda_1)\, \overline{G_{k\ell}}(\lambda_2)\, \Xi(\lambda_1,\lambda_2) \sum_{r} \overline{G^{2}_{rmjr}}(\lambda_1,\lambda_2)\, p^{\vec{1}}_{jm} . \quad (29)$$

The second term breaks up into

$$\sum_{i\ell k} p^{\vec{2}}_{\ell i}\, \overline{G_{ik}}(\lambda_1)\, \overline{G_{k\ell}}(\lambda_2) \sum_{jmr} \Xi(\lambda_1,\lambda_2)\, \overline{G^{2}_{rmjr}}(\lambda_1,\lambda_2)\, p^{\vec{1}}_{jm} \equiv A \cdot B \quad . \quad (30)$$

Let us transform the Wannier sum in A to a Bloch sum. Then since $\langle \vec{k}|p_2|\vec{k}'\rangle = mV_2(\vec{k})\delta_{\vec{k}\vec{k}'}$, and $\bar{G}$ is diagonal in the Bloch representation,

$$A = m \sum_{\vec{k}} V_2(\vec{k})\, \overline{G_{\vec{k}\vec{k}}}(\lambda_1)\, \overline{G_{\vec{k}\vec{k}}}(\lambda_2) = m \sum_{\vec{k}} \frac{\partial \epsilon_{\vec{k}}}{\partial k_2} \frac{1}{(\lambda_1 - \Lambda^*(\lambda_1^+) - \epsilon_{\vec{k}})} \frac{1}{(\lambda_2 - \Lambda^*(\lambda_2^+) - \epsilon_{\vec{k}})} \cdot \quad (31)$$

The propagators are even under inversion ($\vec{k} \to -\vec{k}$) but the velocity $V_2 = \partial\epsilon/\partial k_2$ is odd, giving us zero, and all vertex corrections now vanish.
In this case, the fortunate cancellation of vertex corrections comes about as a consequence of the approximation of the proper self-energy by a site-diagonal quantity, Eq.(12).

Zero-Temperature D.C. Conductivity

We are now in a position to evaluate $\sigma(0)$. Because the vertex corrections vanish,

$$I_{12}(\lambda,\lambda) = \sum_{ij\ell m} p^{\vec{1}}_{jm}\, \overline{\langle m|\delta(\lambda - H)|\ell\rangle}\, p^{\vec{2}}_{\ell i}\, \overline{\langle i|\delta(\lambda - H)|j\rangle}$$

$$= m^2 \sum_{\vec{k}} V_1(\vec{k})\, V_2(\vec{k})\, \overline{\langle \vec{k}|\delta(\lambda - H)|\vec{k}\rangle}^2 . \quad (32)$$

With the definition $\overline{\langle \vec{k}|\delta(\lambda - H)|\vec{k}\rangle} = \frac{-1}{\pi} \operatorname{Im} \overline{G_{\vec{k}\vec{k}}}(\lambda)$, we get for $\vec{1} = \vec{2}$ $(\sigma(\omega) = \sigma_{11}(\omega))$

$$\sigma(0) = \frac{2}{\pi} \int_{-\infty}^{\infty} d\lambda \left[-\frac{\partial f(\lambda)}{\partial \lambda}\right] \sum_{\vec{k}} V_1(\vec{k})^2 \left[\operatorname{Im}\overline{G_{\vec{k}\vec{k}}}(\lambda)\right]^2 \quad , \tag{33}$$

where we have included a factor of two for the two possible spin orientations. At $T = 0$, $-\frac{\partial f(\lambda)}{\partial \lambda} = \delta(\lambda - \mu)$ with $\mu \equiv$ chemical potential, and the conductivity per atom is

$$\sigma(0) = \frac{2}{\pi N} \sum_{\vec{k}} V_1(\vec{k})^2 \left[\operatorname{Im}\overline{G_{\vec{k}\vec{k}}}(\mu)\right]^2 \quad , \tag{34}$$

or

$$\sigma(0) = \frac{2}{\pi} \int_{-\infty}^{\infty} dE \left[\frac{\operatorname{Im}\Lambda^*(\mu^+)}{(\mu - \operatorname{Re}\Lambda^*(\mu^+) - E)^2 + \operatorname{Im}\Lambda^*(\mu^+)^2}\right]^2 \frac{1}{N} \sum_{\vec{k}} V_1(\vec{k})^2 \, \delta(E - \epsilon_{\vec{k}}) \, . \tag{35}$$

This natural separation, only possible for a proper self-energy independent of $\vec{k}$,isolates the lifetime and energy shifts of the single-particle excitations from that part of the conductivity which pertains to the particular lattice under study. We will concentrate on a 3-D simple cubic lattice with $\epsilon_{\vec{k}} = \frac{1}{3}[\cos k_x + \cos k_y + \cos k_z]$. Then, $V_1(\vec{k})^2 = \frac{\sin^2 k_x}{9} = \frac{1}{9}(1 - \cos^2 k_x)$. Consider the functions

$$P_n(E) \equiv \frac{1}{N} \sum_{\vec{k}} \frac{e^{nik_x}}{E - \epsilon_{\vec{k}}} .$$

for which we have

$$\frac{1}{N} \sum_{\vec{k}} V_1(\vec{k})^2 \, \delta(E-\epsilon_{\vec{k}}) = \frac{1}{18\pi} \; \mathrm{Im} \, [P_2(E^+) - P_o(E^+)].$$

The imaginary part of both $P_2(E^+)$ and $P_o(E^+)$ vanish outside the unperturbed band and

$$\sigma(0) = \frac{1}{9\pi^2} \, \mathrm{Im} \int_{-1}^{1} dE \, \left[\frac{\mathrm{Im}\Lambda^*(\mu^+)}{(\mu - \mathrm{Re}\Lambda^*(\mu^+) - E)^2 + \mathrm{Im}\Lambda^*(\mu^+)^2} \right]^2 [P_2(E^+) - P_o(E^+)]. \quad (36)$$

We have calculated the D.C. conductivity for our 3-D alloy in order to illustrate our formal results . Generally, there are two ways in which the D.C. conductivity can vanish. If the density-of-states at the Fermi level is zero then so is $\sigma(0)$. In addition, we can have a finite $\rho(\mu)$, but a zero mobility because of wave function localization. Eq.(36)only admits a zero in $\sigma(0)$ if $\rho(\mu)$ is zero so we cannot take the latter possibility into account . The conductivity is displayed in Figure 11 against its respective density-of-states. There is a rather direct correlation between the magnitude of the density-of-states and that of the conductivity. This relationship is understood in

terms of the availability of states at a given energy to which an initial state can make a transition. We also find peaks in $\sigma(0)$ which we associate with velocity peaks in the cubic band structure.

References

1. R.J. Elliot, J.A. Krumhansl, P.L. Leath, Rev. Mod. Phys. 46, 465 (1974), H. Ehrenreich and L.M. Schwartz, Solid State Phys. 31, 149 (1976), and J.D. Joannopoulos and Marvin L. Cohen, Solid State Phys. 31, 71 (1976).
2. P. Soven, Phys. Rev. 156, 809 (1967), 178, 1136 (1969).
3. D.W. Taylor, Phys. Rev. 156, 1017 (1967).
4. W.H. Butler, Phys. Rev. B 8, 4499 (1973).
5. Pabitra N. Sen and Felix Yndurain, Phys. Rev. B 13, 4387 (1976).
6. R. Friedberg and J.M. Luttinger, Phys. Rev. B 12, 4460 (1975) and references cited therein.
7. B. Velický, S. Kirkpatrick, and H. Ehrenreich, Phys. Rev. B 1, 3250 (1970).
8. For example, while computer time increases expenentially with the number of neighboring shells over which one chooses to perform the exact configurational averages, the use of a Monte Carlo sampling technique may extend our calculations to large numbers of such shells at only a modest increase in computational time.
9. G. Baym and L. Kadanoff, Phys. Rev. 124, 287 (1961).
10. Our Eqs.(6) and (7) were first derived by Vijay Kumar and S.J. Joshi, J. Phys. C: Solid State Phys. 8, L148 (1975), in the context of applying Butler's[4] self-consistent equations to a tetrahedral alloy.
11. A more complete exposition of this entire work is contained in the Ph.D. Thesis of one of us (P.B.).
12. B. Velický, S. Kirkpatrick, and H.Ehrenreich, Phys. Rev. 175, 747(1968).
13. R. Alben, M. Blume, H. Krakauer, and L. Schwartz, Phys. Rev. B 12, 4090 (1975).
14. B. Velický, Phys. Rev. 184, 614 (1969).

Figure Captions

Fig. 1 Real and Imaginary parts of the complex self-energy in the CPA. The two sets of curves are indicative of the type of results that can be expected from this approximation. In part a, we have a situation where the alloy bands overlap while in b the case of split bands.

Fig. 2 Comparison of 1,3, and 5 cluster calculations for $\rho(\omega)$ using $|\text{Im } \Sigma_{\text{trial}}| = .15$ with $V_i = \pm .5$, $c = .5$. Background (histogram) is exact results of Ref. 4.

Fig. 3 Comparison of 1,3, and 5 cluster calculations for $\rho(\omega)$ in 1-D using $|\text{Im } \Sigma_{\text{trial}}| = .5$ with $V_i = \pm 1.0$ and $c = .5$. Background (histogram) is exact results of Ref. 4.

Fig. 4 Curves give $\rho(\omega)$ for 3 different values of Σ_{trial} for a 5 cluster calculation of a 1-D alloy with $V_i = \pm 1.0$ and $c = 5$. The sharpest peaks are associated with lowest value of $|\text{Im } \Sigma_{\text{trial}}|$, .4. Other values are .5 and .6.

Fig. 5 (a) Density-of-states for a 50-50% concentration 1-D alloy with $V_A = 2.0$, $V_B = -2.0$. Histogram is exact calculations from Ref. 4. The full band is obtained by reflecting the portion shown through the origin. These results were computed from Im Σ = -.80 and Re Σ obtained from Im Σ by Eq.(8), including the exact scattering from all configurations of a central atom and its 4 nearest neighbors. (b) Extension of the above results to a cluster of 7 atoms using the same $\Sigma(\omega)$. This result is comparable in accuracy and wealth of detail to the best self-consistent calculation to date, Ref. 4.

Figure Captions

Fig. 6 Curves give $\rho(\omega)$ for 3 different values of Σ_{trial} for a 5 cluster calculation of a 1-D alloy with $V_i = \pm 2.0$ and $c = .5$. The sharpest peaks are associated with lowest value of $|\text{Im } \Sigma_{trial}| = .4$, other values are .8 and 1.0.

Fig. 7 Comparison of 7 cluster calculations of $\rho(\omega)$ (dashed line) using a two step $|\text{Im } \Sigma_{trial}|$ (long dashed line), with numerical work of Alben et. al.[13] (solid line) who solved the Schrodinger equation for an 8000 atom 3-D tight binding solid. The potentials are $V_i = \pm .75$ with impurity concentration of .1. Small horizontal arrows indicate the height to which their peaks rise.

Fig. 8 Comparison of 7 cluster calculation of $\rho(\omega)$ (dashed line) using a 3 step $|\text{Im } \Sigma_{trial}|$ (long dashed line) with results of Alben et. al.[13] (solid line). $V_i = \pm .75$ and $c = .1$ for this 3-D tight binding alloy. The sharp peak in Fig. 7 at $\omega = .86$ is absent because of the coarser energy scale used. Arrow indicates the height to which their peak rises.

Fig. 9 Density-of-states for $V_A = .75$, $V_B = .75$ in a 3-D simple cubic lattice with concentration $c_A = .1$. Cowley short range order parameter α is -.07, .3, and .7 respectively. $|\text{Im } \Sigma|$ is the same as in Fig. 7 (dot-dash line) and $\text{Re}\Sigma$ is obtained therefrom by use of Eq. (8).

Figure Captions

Fig. 10. Solid line is $|\mathrm{Im}\ \Lambda^{*}(\omega)|$ for the 1-D alloy of Figs. 5 and 6 in the 5 cluster approximation while the dashed line is the corresponding $|\mathrm{Im}\ \Sigma_{CPA}(\omega)|$.

Fig. 11 Then D.C. conductivity (dashed line) of a 3-D alloy with $V_i = \pm .75$ and $c = .1$. Density-of-states taken from Fig. 7 is shown in the solid line.

FIG. 1

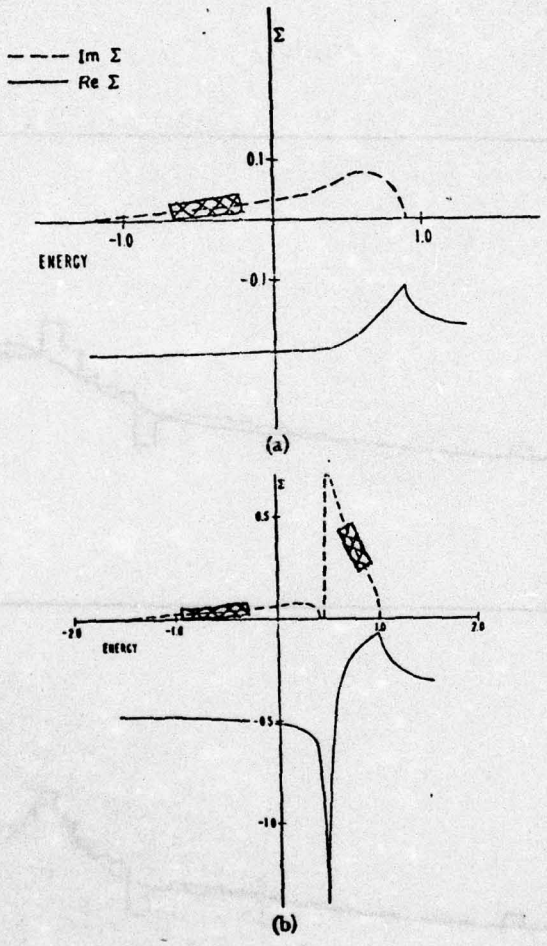


FIG.2

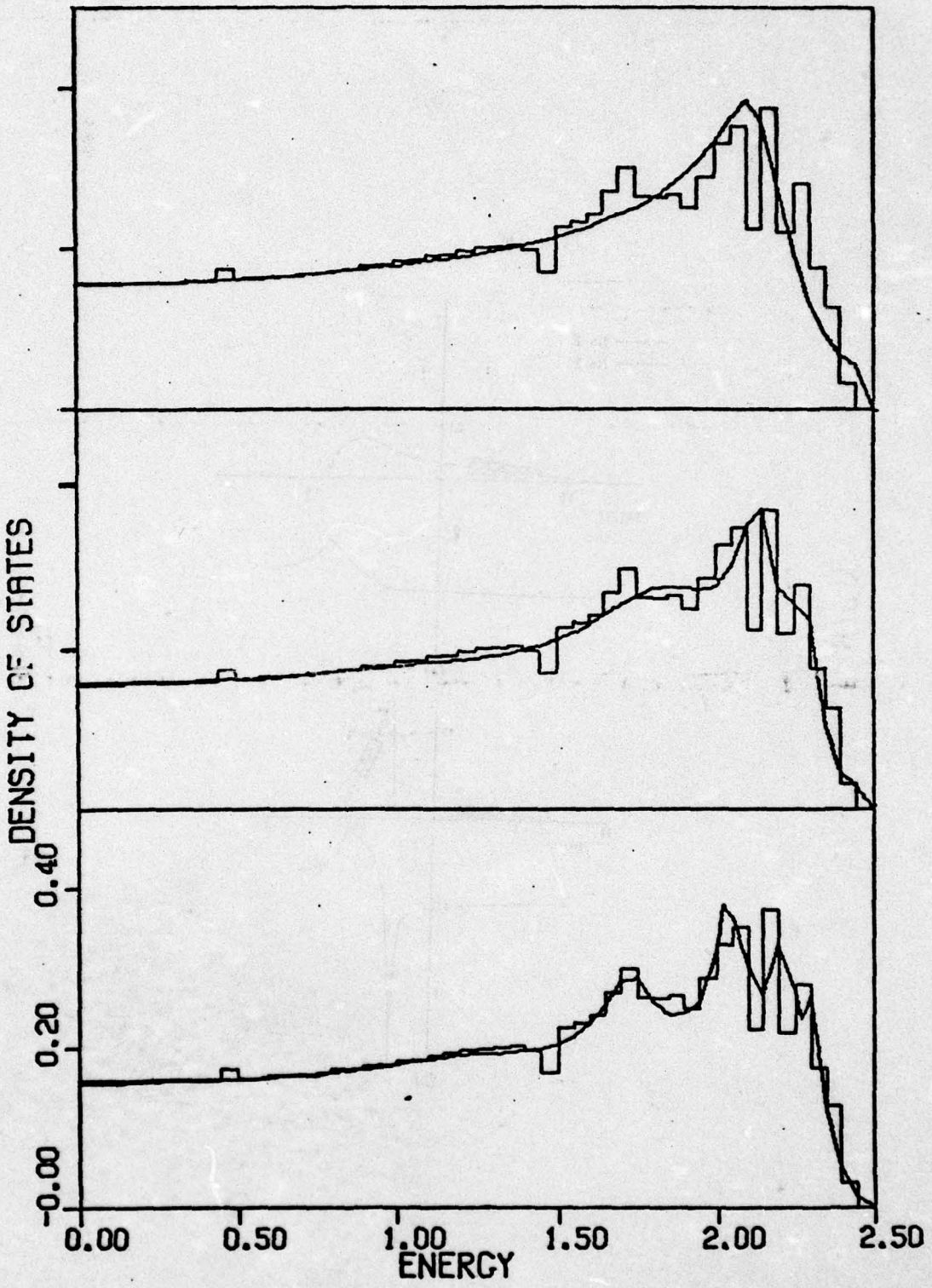


FIG.3

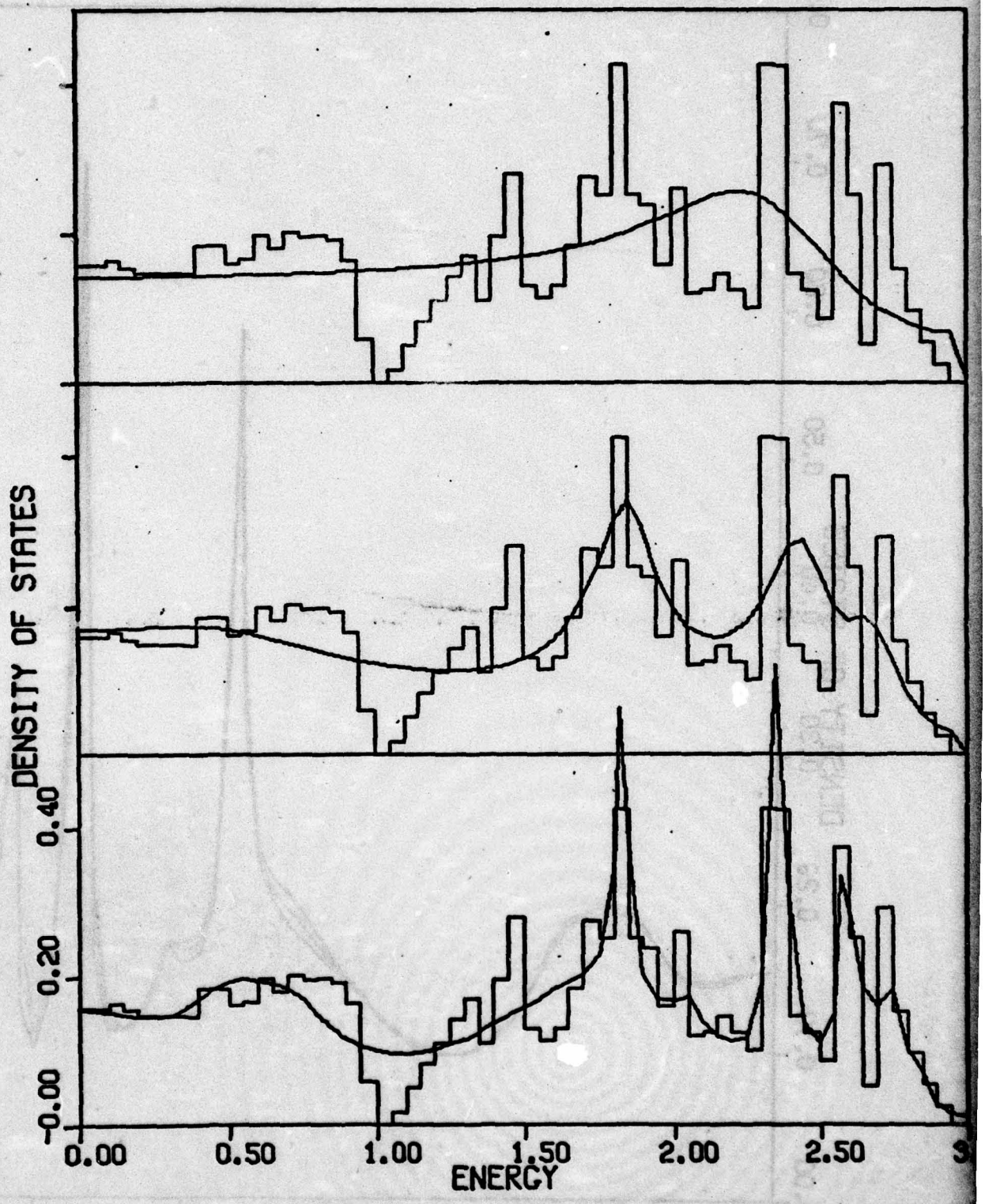
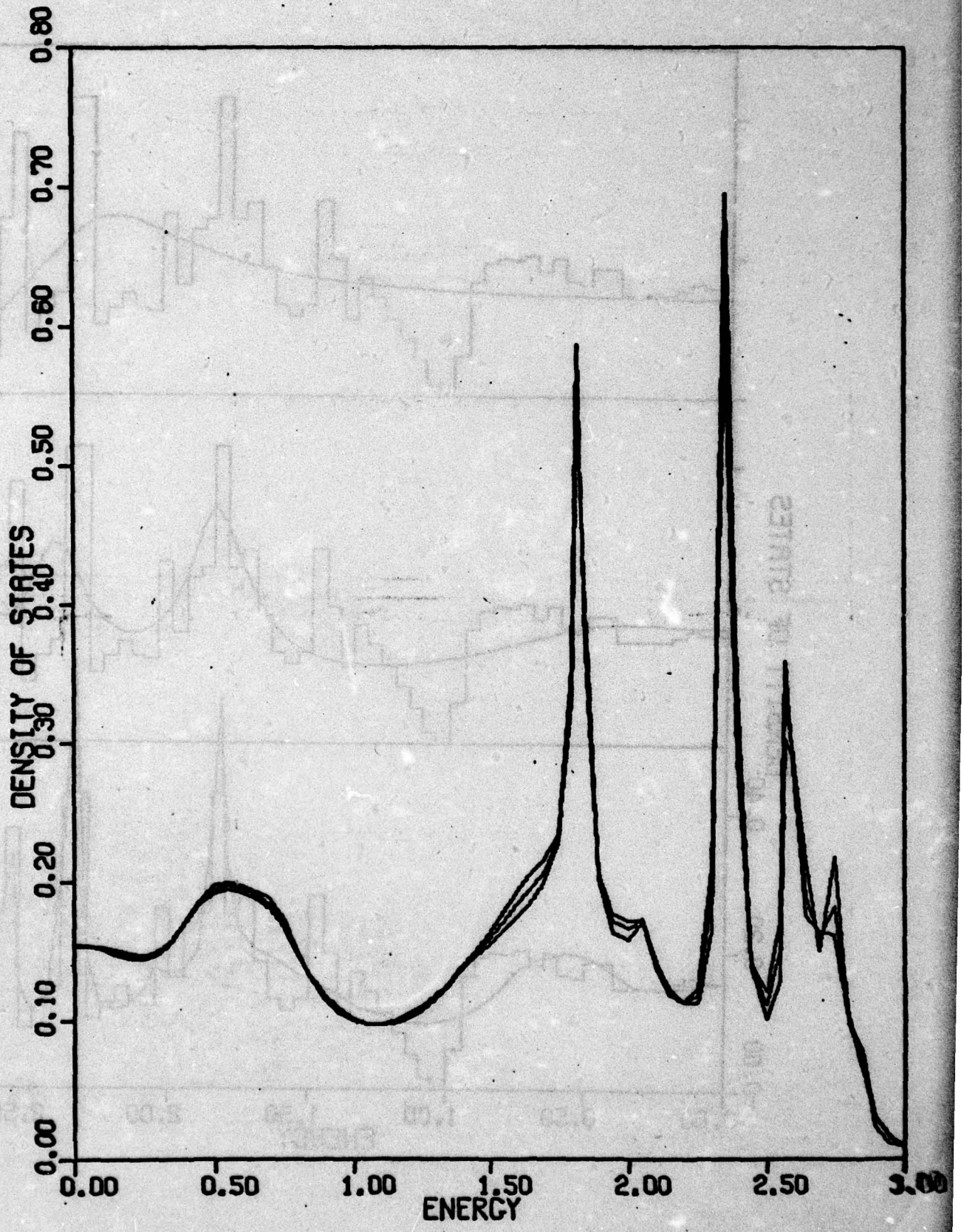


Fig.4


DENSITY OF STATES
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.00
0.50
1.00
1.50
2.00
2.50
3.00
ENERGY

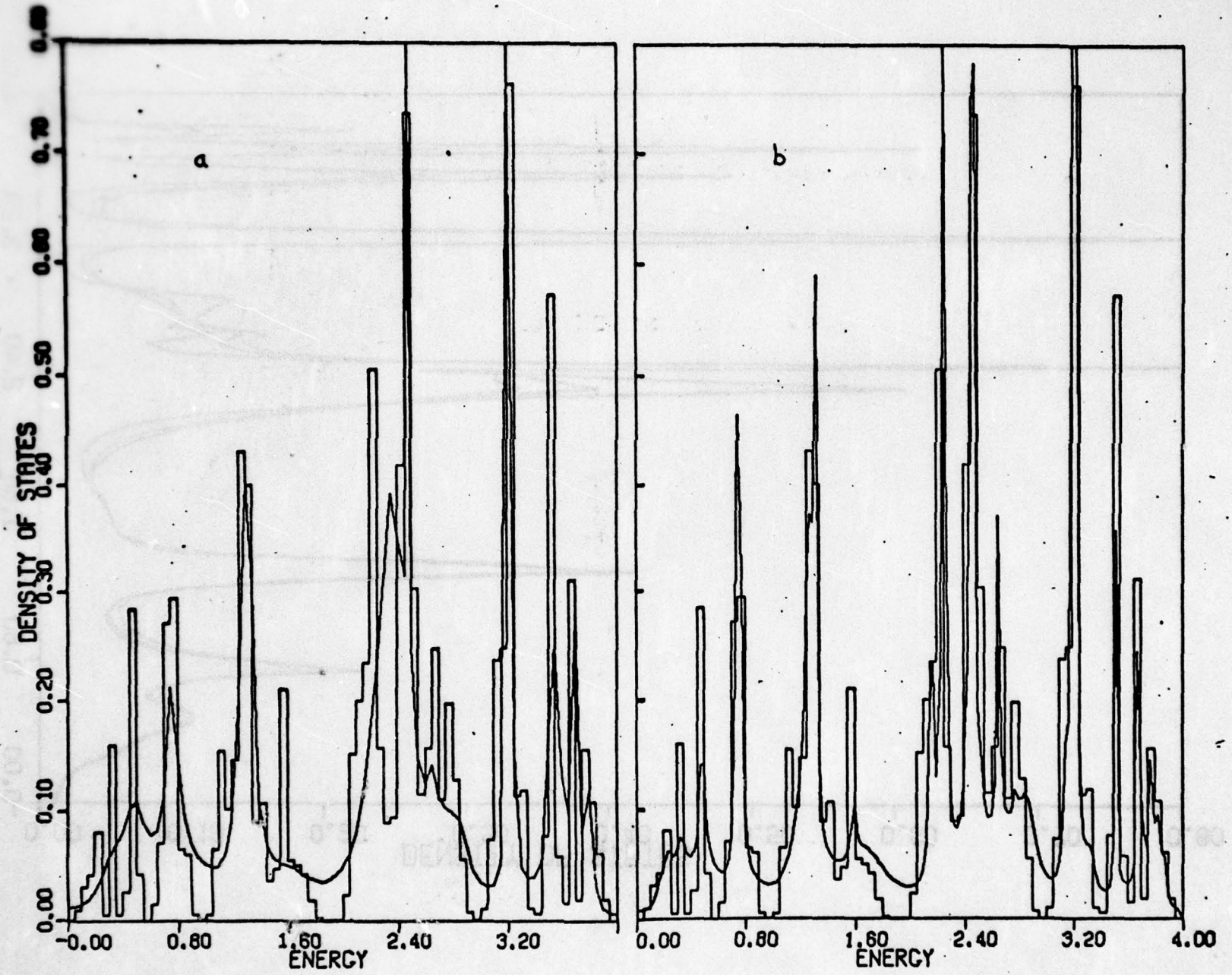

a
b
DENSITY OF STATES
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
-0.00
0.80
1.60
2.40
3.20
4.00
ENERGY
ENERGY

FIG. 6

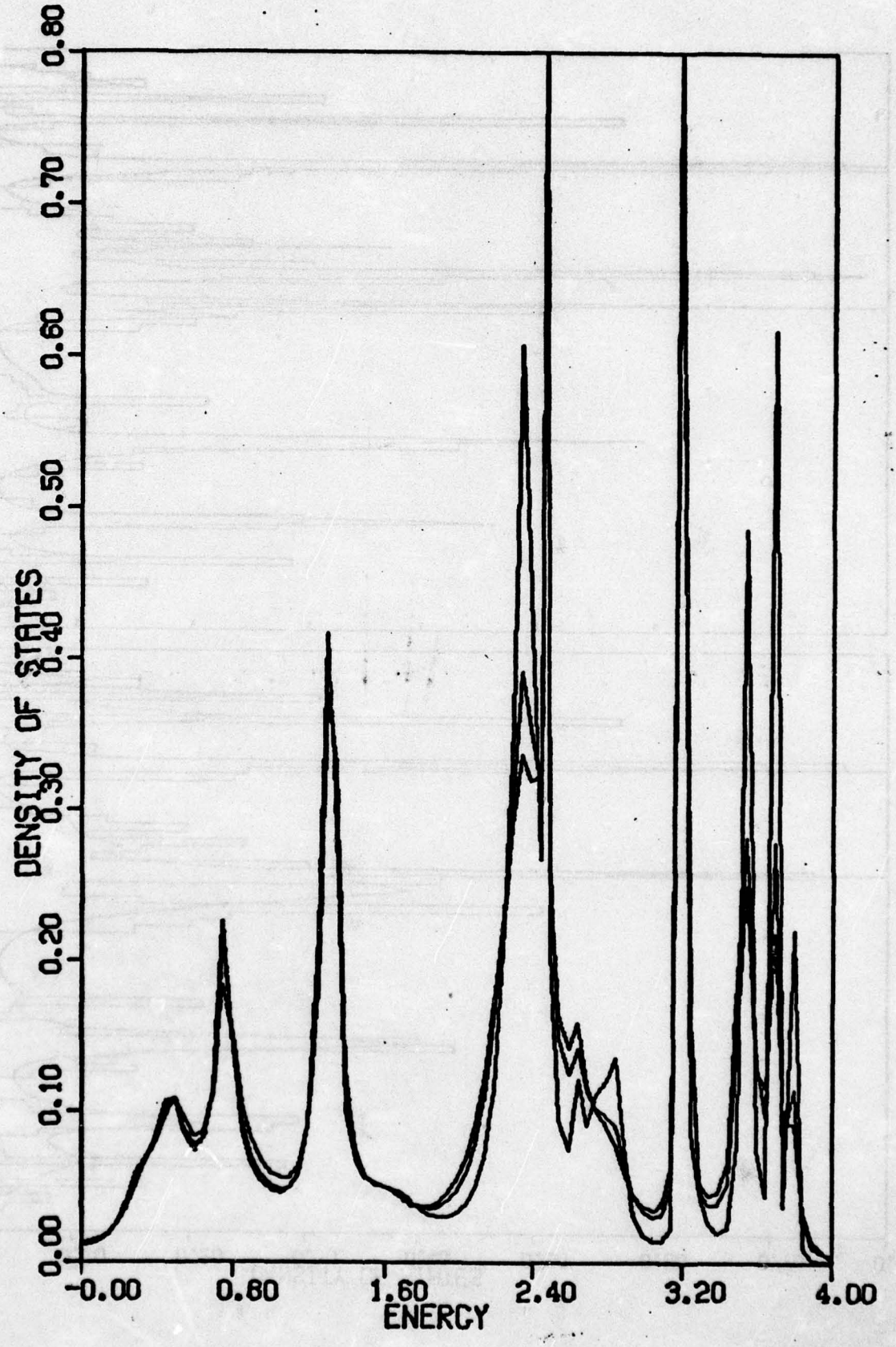

DENSITY OF STATES
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
-0.00
0.80
1.60
2.40
3.20
4.00
ENERGY

FIG.7

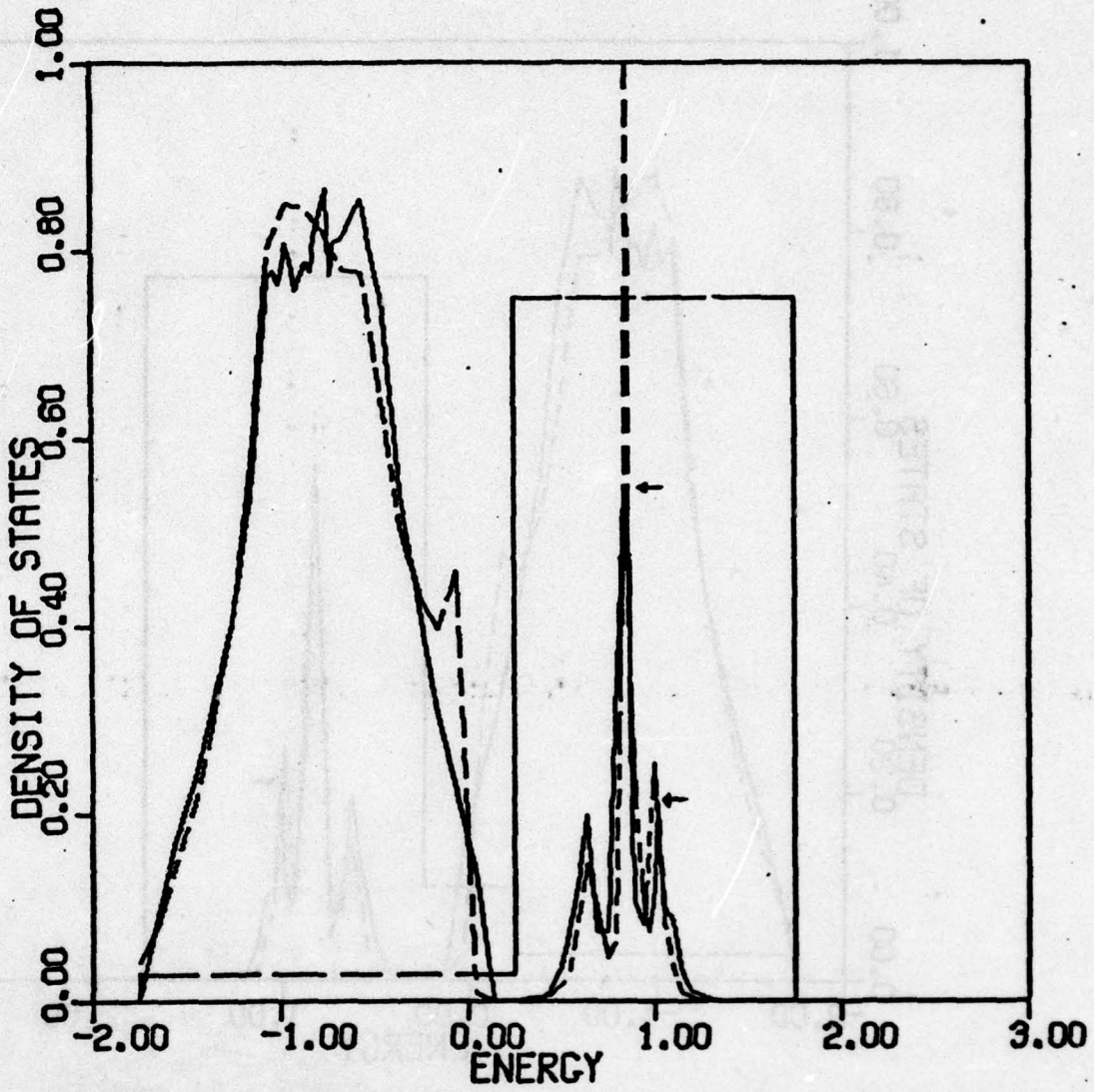
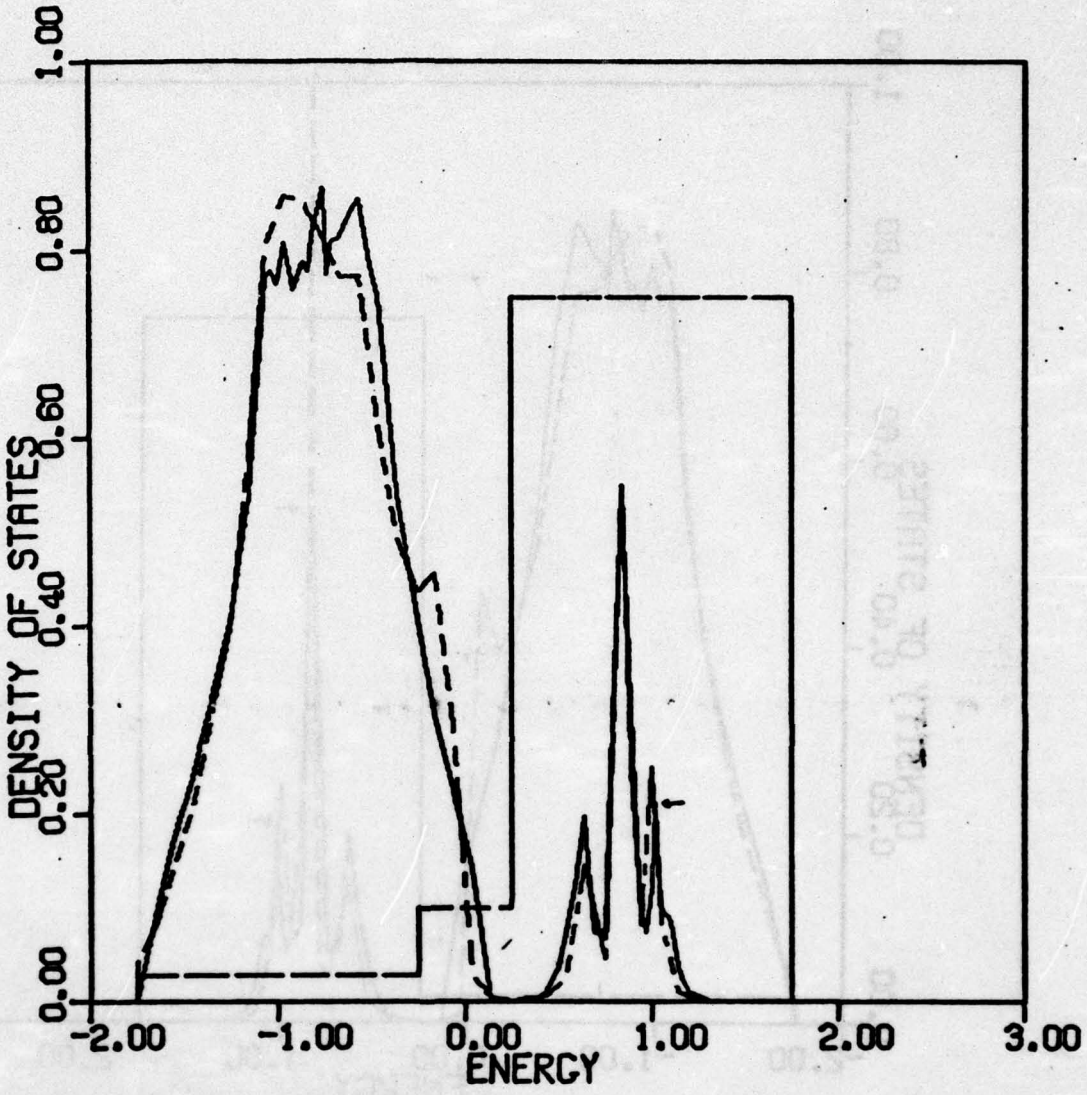

DENSITY OF STATES
0.00
0.20
0.40
0.60
0.80
1.00
-2.00
-1.00
0.00
1.00
2.00
3.00
ENERGY

FIG. 8


DENSITY OF STATES
0.00
0.20
0.40
0.60
0.80
1.00
-2.00
-1.00
0.00
1.00
2.00
3.00
ENERGY

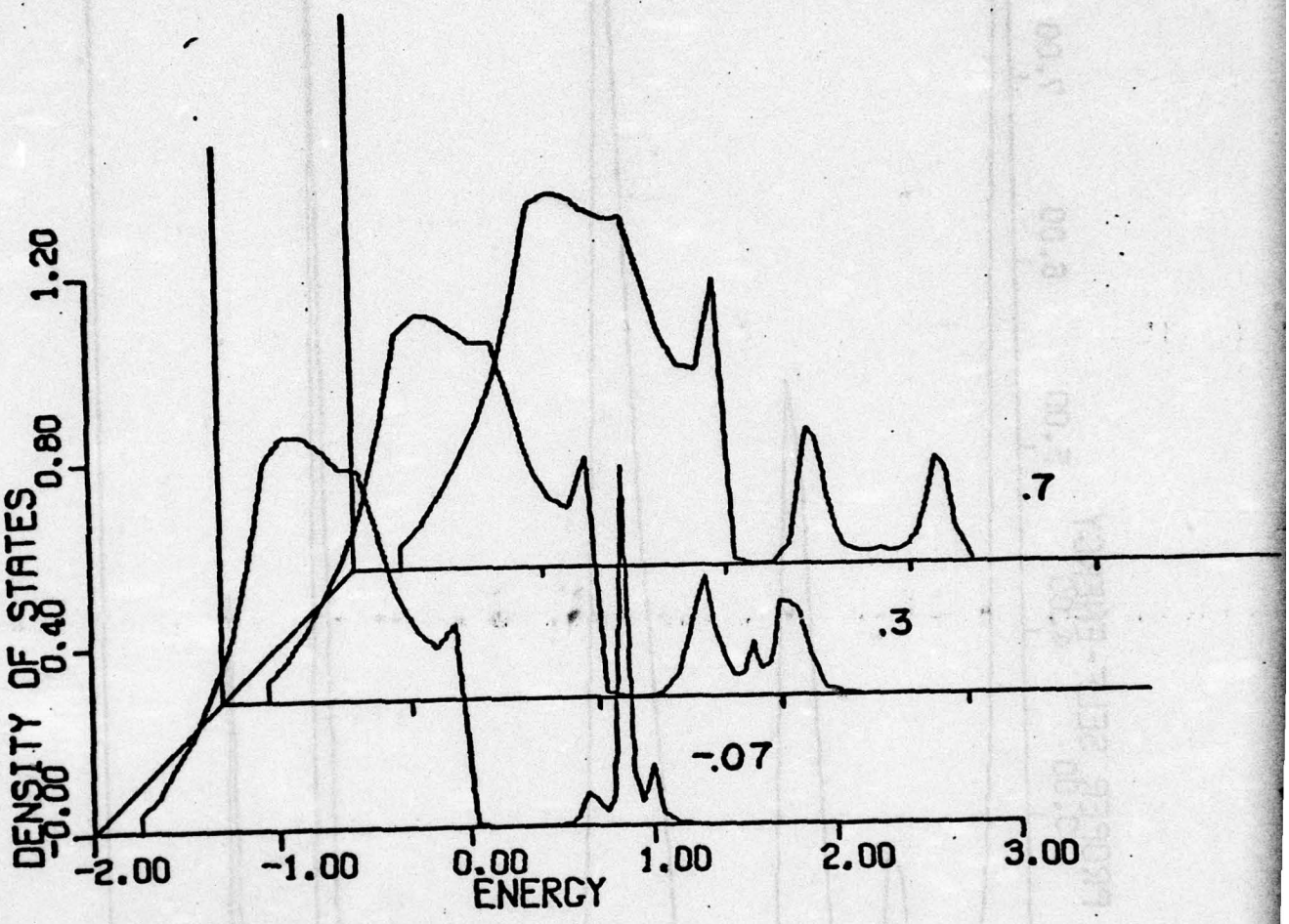
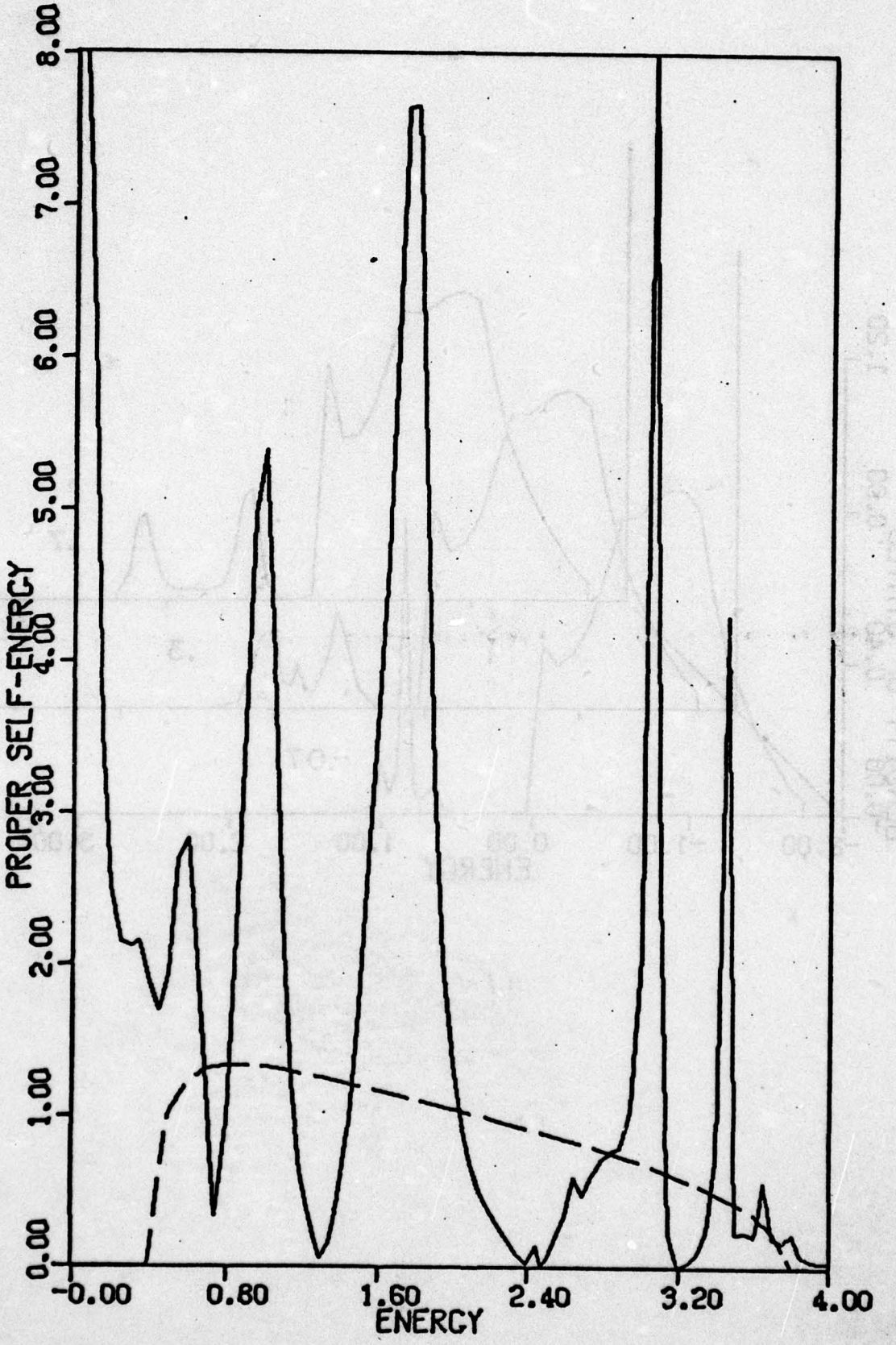
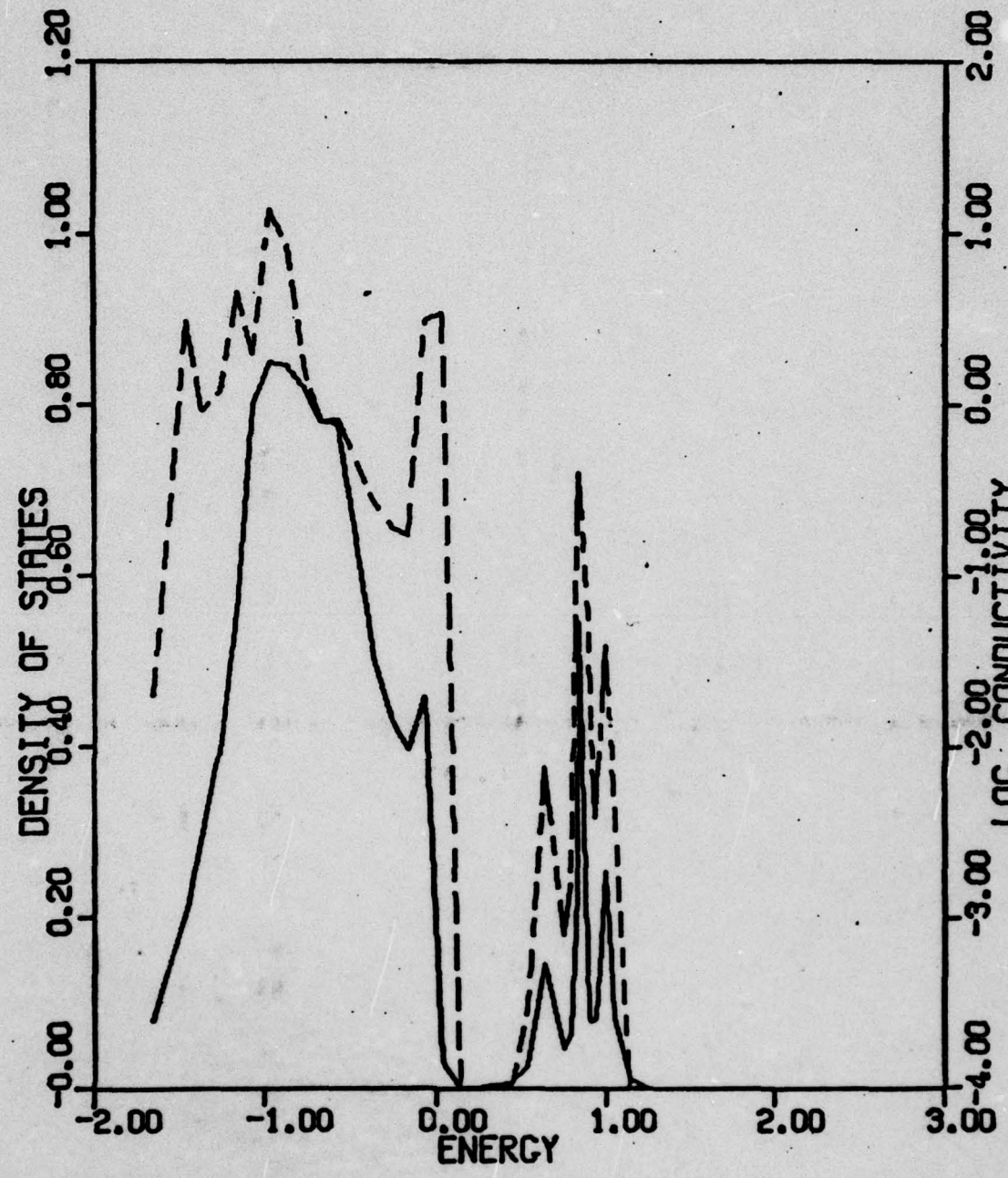

DENSITY OF STATES
0.00
0.40
0.80
1.20
-2.00
-1.00
0.00
1.00
2.00
3.00
ENERGY
.7
.3
-.07

Fig. 10


PROPER SELF-ENERCY
0.00
1.00
2.00
3.00
4.00
5.00
6.00
7.00
8.00
-0.00
0.80
1.60
2.40
3.20
4.00
ENERCY


DENSITY OF STATES
1.20
1.00
0.80
0.60
0.40
0.20
-0.00
LOG 10 CONDUCTIVITY
2.00
1.00
0.00
-1.00
-2.00
-3.00
-4.00
-2.00
-1.00
0.00
1.00
2.00
3.00
ENERGY